Multilocus Sequence typing analyses of *Salmonella enterica* subspecies *enterica*: population structure, asymptomatic carriage and host association

Dissertation
zur Erlangung des akademischen Grades
doctor rerum naturalium
(Dr. rer. nat.)
im Fach Biologie

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät I
der Humboldt-Universität zu Berlin

von
Vartul Sangal (M.Sc. Microbiology)
geboren am 15.08.1977 in Muzaffarnagar, Indien

Präsident der Humboldt-Universität zu Berlin
Prof. Dr. Dr. h.c. Christoph Markschies

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät I
Prof. Dr. Lutz-Helmut Schön

Gutachter/innen: 1. Prof. Dr. Thomas F. Meyer
2. Prof. Dr. Mark Achtman
3. PD Dr. Ulrich Steinhoff

eingereicht am: 13th Oktober 2008
Tag der mündlichen Prüfung: 17th Dezember 2008

Summary

Serovars of *Salmonella enterica* subspecies *enterica* are generally pathogenic to humans and other mammals. In this study, I examined the population structure of one of the most common serovars of this subspecies isolated from humans and food animals, serovar Newport, using a multilocus sequence typing scheme. This scheme was also used to analyze isolates of this subspecies from chronic human carriers and reptiles to determine whether isolates from these sources represent distinct populations than those from other hosts.

Multilocus sequence typing has extensively been used to study evolution and population structure of a wide range of organisms. 400-600 bp fragments of 7 housekeeping genes were sequenced and every unique sequence of each gene fragment was given a distinct allele number. Each unique combination of alleles was assigned a distinct sequence type number. The data were used in further analyses.

Three lineages, namely Newport-I, Newport-II and Newport-III were identified within serovar Newport which were associated to European humans, animals and humans in North America, respectively. Multidrug resistance phenotypes were most common in Newport-II whereas most isolates in Newport-III were pan-susceptible. When compared to other serovars, the numbers of lineages within Newport were higher than for Enteritidis, Kentucky and Typhimurium but lower than for Paratyphi B. Therefore, serovars of *S. enterica* subspecies *enterica* vary greatly in their population structures.

The sequence types observed for isolates from chronic human carriers were generally the most common among human-clinical and animal isolates. Most isolates from non-carrier humans plus animals were genetically identical to the carried isolates within most serovars. Genetic diversity was also comparable between isolates from these sources. These results suggest that salmonellae from chronic human carriers belong to the same population as isolates from non-carrier humans and animals.

For most serovars, most isolates from reptiles were genetically identical to those from humans or other warm blooded animals. However, in serovars Bovismorbificans, Decatur, Miami and Oranienburg, most reptile isolates were genetically distinct from isolates from other hosts. Only few reptile isolates were tested from Bovismorbificans, Decatur and Miami and only few non-reptile isolates were tested from Oranienburg, and in larger numbers of such isolates would be needed to determine whether these differences are statistically significant.

Zusammenfassung

Serovare von *Salmonella enterica* subspecies *enterica* sind im allgemeinen pathogen für Mensch und andere Säugetiere. In dieser Arbeit habe ich anhand eines "Multilocus Sequences Typing" Typisierungsschemas die Populationsstruktur einer der am häufigsten auftretenden Serovaren dieser Subspecies, das aus Menschen und Schlachttieren isolierte Serovar Newport charakterisiert. Dieses Schema wurde auch für die Charakterisierung von Isolaten derselben Subspecies aus humanen Dauerträgern und Reptilien verwandt, um zu bestimmen, ob Isolate aus diesen Quellen sich in ihrer Populationstruktur von denjenigen unterscheiden, die aus anderen Quellen isoliert wurden.

Multilocus Sequences Typing ist eine weitgehend für die Untersuchung der Evolution und Populationsstruktur von einen breiten Spektrum von Organismen verwendete Technik. 400 - 600 bp lange Fragmente von 7 Haushaltsgenen wurden sequenziert, und jede einzelne Sequenz jedes einzelnen Gens wurde eine Allelnummer zugeordnet. Jede einzelne Allelkombination wurde einem Sequenztyp zugeordnet. Die so gewonnenen Daten wurden weiter analysiert.

Drei "Lineages", Newport-I, Newport-II und Newport-III, wurden innerhalb dieses Serovars identifiziert, die jeweils aus Menschen in Europa, Tieren und Menschen in Nordamerika isoliert wurden. Der Multiresistenz-Phänotyp wurde häufiger in Newport II gefunden, während die meisten Newport III Isolate pan-sensitiv waren. Verglichen mit anderen Serovaren war die Anzahl von *"*Lineages*"* innerhalb Newport höher als bei Enteritidis, Kentucky und Typhimurium, aber niedriger als bei Paratyphi B. Das heisst, die Serovare von *S. enterica* subspecies *enterica* variieren stark in ihrer Populationsstruktur.

Die Sequenztypen in Isolaten aus humanen Dauerträgern waren im allgemeinen am häufigsten in Isolaten von klinischen Patienten und Tieren vorhanden. In der Mehrheit der Serovaren waren die meisten Isolate aus Patienten und Tieren genetisch identisch mit solchen, die aus gesunden Trägern isoliert wurden. Die genetische Variabilität war zwischen Isolaten aus diesen Quellen vergleichbar. Diese Ergebnissen deuten daraufhin, dass Salmonellen aus Dauerträgern sowie Isolate aus Patienten und Tieren derselben Population angehören.

Die meisten Serovare aus Reptilienisolaten waren genetisch identisch mit denen von Menschen und warmblütigen Tieren. In den Serovaren Bovismorbificans, Decatur, Miami und Oranienburg hingegen waren die meisten Isolate aus Reptilien genetisch anders als Isolate aus anderen Wirten. Allerdings wurden nur wenige Isolate der Serovaren

Bovismorbificans, Decatur und Miami aus Reptilien und nur wenige Isolate der Serovaren Oranienburg aus anderen Quellen getestet; eine grössere Anzahl von Isolaten müsste daher untersucht werden, um festzustellen ob diese genetischen Unterschiede statistich signifikant sind oder nicht.

Keywords: *Salmonella enterica*, population structure, serovar, Newport, chronic human carrier, reptile, mutation, recombination.

Schlagwörter: *Salmonella enterica*, Populationsstruktur, Serovar, Newport, humanen Dauerträgern, Reptilien, Mutation, Rekombination.

Acknowledgements

My thesis work was carried out at the Department of Molecular Biology, Max-Planck Institute for Infection Biology, Berlin and I would like to thank Prof. Thomas F. Meyer, the Director of the Department of Molecular Biology for granting me the permission to work in the department. I would also like to thank Prof. Martin Vingron, the Managing Director of the Max-Planck Institute for Molecular Genetics, Berlin for financial support during this study.

I would like to express my sincere gratitude to my doctoral supervisor Prof. Mark Achtman for providing me the opportunity to work in his group. I am very grateful to him for his valuable guidance, unwavering support, encouragement and all his efforts to bring out the best in me. It has been a great experience working with him.

I would like to extend my gratitude to Prof. Dr. Richard Lucius for being my thesis supervisor at the Humboldt University, Berlin.

My sincere thanks go to Dr. Giovanna Morelli for her continued help and support. I am greatly indebted to her for her help in arranging my accommodation on my arrival in Berlin and valuable advice in official and personal matters.

I would like to thank Dr. Bodo Linz, Dr. Camila Mazzoni, Dr. Philippe Roumagnac and Dr. Yoshan Moodley for their help in data analyses and insightful discussions. I am thankful to Ms. Steffie Bernhoeft, Mrs. Christiana Stamer, Mrs. Barica Kusecek and Mrs. Mirjam Feldkamp for their assistance in the laboratory.

I would like to thank Dr. Heather Harbottle, U. S. Food and Drug Administration, Laurel, U.S.A. for providing me MLST data and the antimicrobial resistance typing information for 135 *Salmonella* isolates. I thank Dr. Reiner Helmuth and Dr. Beatriz Guerra Roman from the Federal Institute for Risk Assessment (BfR), Berlin for providing *Salmonella* isolates from reptiles and their help with serotyping and antimicrobial resistance typing. I am grateful to Dr. Wolfgang Rabsch and Dr. Angelika Fruth from the Robert-Koch Institute, Wernigerode, Germany; Dr. François-Xavier Weill, Institute Pasteur, Paris; Dr. Fidelma

Boyd, University College Cork, Ireland; Dr. Roy Curtiss III, Arizona State University, Tempe, U.S.A.; Dr. Mia Torpdahl, Statens Serum Institute, Copenhagen, Denmark; Dr. Helene Andrews-Polymenis, College of Medicine, College Station, Texas, U.S.A.; and Dr. Steffen Porwollik, Sidney Kimmel Cancer Center, San Diego, U.S.A. for providing *Salmonella* isolates.

I am thankful to Prof. Lothar H. Wieler and Dr. Antina Lübke-Becker from the Veterinary Faculty of the Freie University, Berlin for teaching me the methods of primary identification of *Salmonella enterica* isolates. I am grateful to Prof. Martin C. J. Maiden and Dr. Keith Jolley from the Department of Zoology of the University of Oxford, Oxford for inviting me to their laboratory to learn DNA sequencing.

I thank Mr. Ralf Traeger for IT support and the library staff of the Max-Planck Institute for Infection Biology for efficient help in getting me re-prints. I am also thankful to Mr. Vimalkumar Velayudhan, Environmental Research Institute, UCC, Cork, Ireland for IT support.

Abbreviations

AFLP	amplified fragment length polymorphism
bp	base pair (s)
cm	centimetre
DLV	double locus variant
DNA	deoxyribonucleic acid
dNTPs	deoxynucleoside triphosphates
ds	double stranded
EDTA	ethylene-diamine-tetraacetic acid
ET	enzyme type
ExoI	exonuclease-I
g	gram
g'	gravitational acceleration (g' = 9.81 m/s²)
h	hour
LB medium	Luria-Bertani medium
LPS	lipopolysaccharides
µl	microlitre
M	molar
MDR	multidrug resistance
mg	milligram
min	minutes
ml	millilitre
MLEE	multilocus enzyme electrophoresis
MLST	multilocus sequence typing
MLVA	multilocus variable number of tandem repeats analysis
mM	milli-molar
MS_{TREE}	minimal spanning tree
n.d.	not done
ncHA	non-carrier humans plus animals
NhWBA	non-human warm blooded animals
OD	optical density

PBS	phosphate buffered saline
PCR	polymerase chain reaction
PFGE	pulse-field gel electrophoresis
pmol	picomoles
RAPD	random amplified polymorphic DNA
RNA	ribonucleic acid
RNase	ribonuclease
rpm	revolutions per minute
rRNA	ribosomal RNA
SAP	shrimp alkaline phosphatase
SDR	single drug resistance
SID	Simpson's index of diversity
SLV	single locus variant
SNP	single nucleotide polymorphism
ST	sequence type
TAE	Tris-acetate-EDTA
TLV	triple locus variant
Tris	Tris (hydroxymethyl) aminomethane
U	unit (s)
V/cm	volts per centimeter
VNTR	variable number of tandem repeats
w/v	weight/volume

Contents

1. Introduction

Salmonellae are facultatively anaerobic, gram-negative rod shaped bacteria that belong to the family Enterobacteriacae (Chart, 2003). They are one of the major causes of morbidity among humans (Herikstad *et al.*, 2002; Mead *et al.*, 1999). *Salmonella* infection in humans may result in self limiting diarrhoea to life threatening systemic diseases (Chart, 2003). Based on the description of the clinical symptoms, typhoid fever and dysentery found to be responsible for the mass deaths in Jamestown, Virginia during the English colonization between 1607 and 1624 (Earle, 1979). Alexander the Great (356-323 BC) and Prince Albert the Consort of Queen Victoria (1819-1861), two famous historical figures, also died of typhoid fever[1] (Knights, 1969; Cunha, 2004). The typhoid epidemic has been responsible for high mortality rate among the American soldiers during the Spanish-American war in 1898 (Cirillo, 2000). The bacilli were first isolated in 1885 from pigs and named after Daniel Elmer Salmon, one of the discoverer (Clark, 1959).

1.1 Classification and Nomenclature

Salmonella is a member of the family Enterobacteriaceae of the class γ- proteobacteria in the kingdom Bacteria (Euzéby, 2007).

Kingdom	Bacteria
Phylum	Proteobacteria
Class	γ-proteobacteria
Order	Enterobacteriales
Family	Enterobacteriaceae
Genus	*Salmonella*

The nomenclature of *Salmonella* had been very unstable and undergone dramatic changes in the last few decades (Euzeby, 1999; Brenner *et al.*, 2000; Ezaki *et al.*, 2000a; Ezaki *et al.*, 2000b; Yabuuchi and Ezaki, 2000; Heyndrickx *et al.*, 2005; Tindall *et al.*, 2005).

[1]Various studies argued for other diseases to be the cause of deaths of Alexander the Great as well as Prince Albert because similar symptoms are also observed for various non-typhoidal diseases (Marr and Calisher, 2003; Williams and Arnott, 2004; Paulley, 1993).

S. choleraesuis, S. typhi and *S. enteritidis* had been considered three species within the genus *Salmonella* until 1973 when results from DNA-DNA hybridization experiments led to the proposal that *S. choleraesuis* be the type and only *Salmonella* species (Crosa *et al.*, 1973). Subsequently, in 1999, Euzeby recommended changing this name to *S. enterica* and subdividing the species into six subspecies namely *enterica*, *salamae*, *arizonae*, *diarizonae*, *houtenae* and *indica* (Euzeby, 1999). These subspecies are also designated by roman numbers I, II, IIIa, IIIb, IV and VI, respectively. The proposal to grant the status of a species to *S. bongori* (Reeves *et al.*, 1989) which was earlier considered as subspecies V, was also accepted (Euzeby, 1999).

In 2004, *S. subterranea* was published as a new species, isolated from a low pH subsurface sediment contaminated with nitrite and Uranium-VI (Shelobolina *et al.*, 2004). This isolate showed 96.4% homology with *S. bongori* in 16S ribosomal DNA sequences. Currently, the genus *Salmonella* consists of three species namely *S. enterica, S. bongori* and *S. subterranea* and *S. enterica* is further divided into six subspecies.

1.2 Classical schemes for *Salmonella* typing

Two classical methods are routinely used for subtyping of *S. enterica*, namely serotyping and phage-typing.

1.2.1 Serotyping

Serotyping scheme was proposed by Kauffmann and White in 1957 for antigenic classification of *S. enterica*. The outer membrane of Gram-negative bacteria including *S. enterica* contain two sets of immunogenic antigens, one being long-chain lipopolysaccharides (LPS), also referred to as somatic or O antigens and the second being flagellar protein subunits, also called H antigens (Chart, 2003). In addition, the surface polysaccharide antigen Vi (virulence antigen) is expressed by serovars Typhi, Paratyphi C and occasionally by serovar Dublin (Selander *et al.*, 1990b; Selander *et al.*, 1992; Cheminay and Hensel, 2008). LPS molecules consists of three regions: (i) a hydrophobic lipid A residue, (ii) the core oligosaccharide region and (iii) the hydrophilic O-polysaccharide or O-chain region (Raetz and Whitfield, 2002). The O-chain region consists of 20-40 repeating units of 2-8 monosaccharides. Differences in sugar composition and the degree of branching are the source of variation among O-antigens (Raetz and Whitfield, 2002).

Most of *S. enterica* isolates are equipped with two genes, namely *fliC* and *fljB*, that encode H1 (Phase 1) and H2 (Phase 2) flagellar antigens, respectively (Iino *et al.*, 1988). These genes were formerly known as *H1* and *H2*, respectively. Only one of these genes is expressed at any one time and changes in expression are regulated by a phase switch (Silverman *et al.*, 1979; Simon *et al.*, 1980). This switch is an invertible DNA sequence adjacent to the *fljB* gene which contains a promoter region required for transcription. In one orientation *fljB* is co-expressed with *fljA* (formerly known as *rh1*) whose product represses the expression of *fliC*. In the other orientation, only *fliC* is expressed (Simon *et al.*, 1980). However, both antigens can be simultaneously expressed by a bacterial culture where a fraction of cells expresses one antigen and the rest express the other antigen.

S. enterica cultures are tested against a panel of antisera to identify O and H antigens and the agglutination profiles are used to assign the isolates to serovars. A cell suspension is mixed and incubated with an antiserum and a reaction is scored positive on the formation of visible clumps (The National Food Institute, 2008). Bacterial cultures are grown on a non-selective agar medium to test O antigens and semisolid agar medium to test H antigens. Various kits containing omni-, poly- and mono-valent antisera are commercially available. Each unique combination of O and H antigens (antigenic formula) is assigned a serovar designation which is usually the name of the geographic location where the strain was first isolated (Brenner *et al.*, 2000). However, naming of the serovars for subspecies other than *enterica* is no longer practised. O-antigens are denoted by numbers, H1 antigens by lower case letters and H2 by both number and letters (Popoff, 2001). The three types of antigens are separated by colons while multiple antigenic determinants within an antigen type are separated by commas. *Salmonella* serovars containing common somatic antigens are often assigned to a sero-group or O group, e.g., serovar Typhimurium belongs to sero-group O:4 that contain all the serovars with O-antigens 4 and 12 (Popoff, 2001).

Example: the antigenic formula for *S. enterica* subsp. *enterica* (henceforth referred to as subspecies *enterica*) serovar Typhimurium is 1,4,[5],12:i:1,2. An underlined somatic antigen indicates that the determinant is caused by the lysogenization of a bacteriophage and factors in square brackets indicate that the expression of these determinants is variable. Some antigenic determinants are weakly agglutinable and are written in brackets '( )' (Popoff, 2001). 1,4,[5],12 in this formula are the somatic antigens that are expressed by isolates of serovar Typhimurium while i and 1,2 are the H1 and H2 antigens, respectively. This scheme has identified more than 2500 serovars to date, 1500 in subspecies *enterica* and remainder in other *S. enterica* subspecies plus *S. bongori* (Popoff *et al.*, 2004).

1.2.2 Phage typing

Phage typing has been practised for identification and classification of *Salmonella* from 1920s, prior to the introduction of serotyping and was extensively used for subtyping of typhoid and paratyphoid bacilli (Felix, 1951; Gershman, 1976). It has also been used for typing of other serovars such as Anatum (Gershman, 1974), Enteritidis (Gershman, 1976), Heidelberg and Typhimurium (Ibrahim, 1969). Phage typing is based on the selective ability of bacteriophages to infect certain *S. enterica* strains and isolates within a serovar are differentiated into various subtypes based on the patterns of their lysis by a series of bacteriophages (Schmieger, 1999). The susceptibility of a bacterial isolate to particular bacteriophages depends on the presence of appropriate phage receptors on the cell surface. The numbers, sizes and transparencies of plaques are compared to known phage patterns to define the phage type (Hickman-Brenner *et al.*, 1991). *Salmonella* isolates that do not react to any of the typing phages are called non-typeable (NT) and if a lysis pattern does not correspond to any of the recognized phage types, it is defined as RDNC (React but does not conform).

Phage typing has been used for routine strain characterization and epidemiological investigations for various *S. enterica* serovars. Ten of the Typhi outbreaks were characterized to have caused by phage type E1 between 1960-1999 in the U.S.A. (Olsen *et al.*, 2003). Host restricted subtypes of certain serovars have also been identified by phage typing, e.g., Typhimurium phage types DT2 and DT99 are restricted to pigeons (Rabsch *et al.*, 2002). Phage typing has also been used to differentiate salmonellae of various sero-groups (Gershman and Markowsky, 1983).

1.3 Host range

Salmonellae are capable to infect a variety of hosts ranging from cold-blooded reptiles through warm-blooded animals and humans. Most serovars of subspecies *enterica* can also infect a wide spectrum of hosts and are referred to as unrestricted or generalist salmonellae, e.g. Tyhimurium and Enteritidis (Uzzau *et al.*, 2000). In contrast, serovars that are specific for particular hosts and are only rarely isolated from other sources are called host-adapted salmonellae (Kingsley and Baumler, 2000). Serovars Choleraesuis and Dublin are adapted to pigs and cattle, respectively, and human cases of Choleraesuis and Dublin infections are often linked to the consumption of contaminated meat products (Kingsley and Baumler,

2000). Still other serovars are exclusively associated with a single host and are called host-restricted serovars, e.g., Typhi can infect only humans and higher primates (Kingsley and Baumler, 2000; Uzzau *et al.*, 2000).

Numerous insertions, deletions and genomic rearrangements have been identified in host-adapted and host-restricted serovars (Wu *et al.*, 2005; Helm *et al.*, 2004). Genomic rearrangements at *rrn* operons have been observed in Typhi (human-restricted) (Liu *et al.*, 2006) and Typhimurium phage type DT2 (pigeon-restricted) (Helm *et al.*, 2004). Various insertions, deletions and genomic rearrangements have been found in Gallinarum and Pullorum, two fowl-restricted variants of serovar Enteritidis (Liu *et al.*, 2002; Wu *et al.*, 2005). Several genomic insertions have been observed in the human-restricted serovars Typhi and Paratyphi A that are usually absent in other serovars and might potentially contribute to the increased virulence of these serovars to humans (McClelland *et al.*, 2004). Host adaptation in Typhi and Paratyphi A has also been accompanied by genome degradation of regions that are apparently not required for infection and survival in humans (McClelland *et al.*, 2004). Some host-adapted and host-restricted serovars are auxotrophs, e.g., Typhi, Gallinarum, Paratyphi A, and depend on their hosts for specific nutrients (amino acids and vitamins) (Uzzau *et al.*, 2000). The auxotrophies may reflect genomic rearrangements that have resulted in the loss of function of genes involved in the biosynthesis of particular nutrients.

1.4 Asymptomatic carriers

Non-typhoidal *Salmonella* infection usually results in self limiting diarrhoea (Chart, 2003). After clinical recovery, patients often excrete salmonellae for 12 days to 5 weeks (Sirinavin *et al.*, 2004). Approximately 1% of such patients become chronic asymptomatic carriers and excrete salmonellae for years (D'Aoust, 1991). The frequency of chronic asymptomatic carriers after infection of serovar Typhi is slightly higher and 1-5% of infections result in long-term colonization of the bacilli in the hepatobiliary system (Gupta *et al.*, 2006). “Mr. N the milker” and Mary Mallon (“Typhoid Mary”) are two widely known chronic human carriers of serovar Typhi (Hasian, 2000; Mortimer, 1999). In addition to infect the healthy population, chronic typhoid and paratyphoid carriers themselves are at a high risk of developing cancer of the gall bladder (Caygill *et al.*, 1994). Therefore, a treatment with antibiotics or cholecystectomy has been recommended for them, both for eradication of

chronic carriage as well as to reduce the cancer mortality in chronic carriers (Caygill *et al.*, 1994).

Salmonella not only persist asymptomatically among humans but also in reptiles and other warm-blooded animals, especially dairy and food animals (Cobbold *et al.*, 2006; McCain and Powell, 1990; Sadeyen *et al.*, 2006; Ebani *et al.*, 2005). They colonize lymph nodes and internal organs and are shed continuously or intermittently with the faeces (McCain and Powell, 1990; Nielsen *et al.*, 2004; Perron *et al.*, 2007). A reservoir of *Salmonella* is maintained in the caeca by avian *Salmonella* carriers (Duchet-Suchaux *et al.*, 1995). However, transient colonization of salmonellae has also been observed in avian liver and spleen (Duchet-Suchaux *et al.*, 1995). Reptiles are probably the largest reservoir of all *Salmonella* subspecies (Briones *et al.*, 2004; Ebani *et al.*, 2005), where they colonize the intestinal tract as commensals (Baumler *et al.*, 1998) and are regularly shed with the excrement (Geue and Loschner, 2002; Briones *et al.*, 2004; Ebani *et al.*, 2005).

1.5 Epidemiology

Salmonellosis claims thousands of lives every year and is a global problem (World Health Organization, 2005). Despite advanced medical science and improved hygiene, *Salmonella* constitutes a major threat to humans and animals, both in developing and developed countries. Approximately 1.4 million human cases of nontyphoidal salmonellosis occur annually in the U.S.A. resulting in 400 deaths (Voetsch *et al.*, 2004), and approximately 200,000 human infections claim ~140 lives per year in Europe (van Lier and Havelaar, 2007). Infants, children, pregnant women, elderly people and immunocompromised patients are at high risk (Mermin *et al.*, 1997; Hemsworth and Pizer, 2006; Milstone *et al.*, 2006; Pasmans *et al.*, 2007). The faecal-oral route is the primary means of *Salmonella* transmission (Herikstad *et al.*, 2002; Mead *et al.*, 1999). Transmission is one of the major challenges for the animal and food processing industries because outbreaks in humans are often linked to the consumption of contaminated meat and food products (World Health Organization, 2005; Voetsch *et al.*, 2004). Some of the outbreaks of particular serovars of subspecies *enterica* that occurred in the last two years are listed in Table 1.1.

Table 1.1 *Salmonella* outbreaks occurred in 2006 and 2007.

Serovar	Country	Year	Source	Reference
Paratyphi B var Java	Sweden, Denmark, U.K.	2007	Baby spinach[1]	(Denny *et al.*, 2007)
Senftenberg	U.K.	2007	Fresh basil	(Pezzoli *et al.*, 2007)
Stanley	Sweden	2007	Alfalfa sprouts	(Werner *et al.*, 2007)
Typhimurium	U.S.A.	2007	Milk and cheese	(Centers for Disease Control and Prevention, 2007b)
Weltevreden	Scandinavia	2007	Alfalfa sprouts	(Emberland *et al.*, 2007)
Othmarschen	South Korea	2007	Catered food	(Kim *et al.*, 2007)
Enteritidis	Hungary	2007	Cooked food	(Krisztalovics *et al.*, 2007)
Kedougou	Norway	2006	Salami	(Emberland *et al.*, 2006)
Bareilly and Virchow	Sweden	2006	Cooked food	(de Jong *et al.*, 2007)
Enteritidis phage type 13a	United Kingdom	2006	Eggs	(Morgan *et al.*, 2007)
monophasic 4,[5],12:i:-	Luxembourg	2006	Pork[1]	(Mossong *et al.*, 2007)
Oranienburg	U.S.A. and Canada	2006	Fruit salad	(Centers for Disease Control and Prevention, 2007a)
Kottbus	Spain	2006	Bottled water	(Palmera-Suarez *et al.*, 2007)
Newport, Braenderup and Typhimurium	U.S.A.	2005-2006	Tomatoes	(Centers for Disease control and Prevention, 2007)

[1]Source suspected but was not confirmed

Although salmonellosis in humans is often associated with food, numerous cases of reptile-associated salmonellosis have also been reported in Europe (Geue and Loschner, 2002; Willis *et al.*, 2002), Asia (Mahajan *et al.*, 2003; Nakadai *et al.*, 2005), the U.S.A. (Centers for Disease Control and Prevention, 2007c), Australia (Thomas *et al.*, 2001) and Canada (Woodward *et al.*, 1997). Of the 1.4 million annual cases of *Salmonella* infection in the United States, 6% are acquired from reptiles (Mermin *et al.*, 2004; Voetsch *et al.*, 2004).

A large proportion of *Salmonella* infections in humans are caused by serovars belonging to subspecies *enterica*. In the U.S.A. in 2005, the top ten serovars that infected humans were Typhimurium, Enteritidis, Newport, Heidelberg, Javiana, 4,[5],12:i:-, Montevideo, Muenchen, Saintpaul and Braenderup, all of which belong to subspecies *enterica* (Centers for Disease Control and Prevention, 2006). Enteritidis, Typhimurium, Virchow, Hadar, Infantis, Newport, Saintpaul, Typhi, Agona and Stanley were the most prevalent serovars in Europe in the years 2000-2006 (European Centre for Disease Prevention and Control, 2008) and again all belong to subspecies *enterica*.

1.6 Other typing schemes for *Salmonella*

Serotyping has widely been used for strain identification and routine diagnostics within *Salmonella* but the technique has limitations. The LPS antigens of some of *Salmonella*

serovars are cross-reactive with other genus of Enterobacteriaceae, especially *Escherichia coli*, which can lead to false results (Rundlof *et al.*, 1998; Navarro *et al.*, 2007). Serotyping groups isolates based on antigenic profiles and does not discriminate within serovars between host-restricted strains and others, e.g., within Typhimurium phage types DT2 and DT99 are avian-restricted whereas DT104 and DT204 are of broad host range.

Phage typing can discriminate between closely related strains but its discriminatory abilities are limited by the available typing phages and some isolates may not be typeable with this technique. Phage typing requires the maintenance of multiple bacteriophages that are usually accessible only to reference laboratories (Foley *et al.*, 2007). Furthermore, the interpretations of the results require considerable experience. One important aspect of phage typing is the long-term reliability of the bacteriophages stocks. Yet, variation among archived phage cultures has been observed over time that could possibly be due to host specific modifications or recombination with residing prophages in the propagation strains (Schmieger, 1999; Rabsch *et al.*, 2004). The phage type of individual strains can also change over time possibly due to the loss of prophages, recombination within prophages, mutations leading to changes in phage receptors at the cell surface and mutations in restriction-modification systems (Rabsch *et al.*, 2004).

An ideal typing method should be user friendly, easy to interpret, reproducible, economic and portable between laboratories. Both serotyping and phage typing are time consuming, require considerable technical expertise for interpretation and sometimes show poor reproducibility between different laboratories. Various other techniques have been used for *Salmonella* typing to overcome some of these problems, some of which are briefly described below:

1.6.1 Antimicrobial susceptibility typing

Antimicrobial therapy is often used to treat systemic salmonellosis. Fluoroquinolones have been commonly used to treat adults and third generation cephalosporins for children (Dunne *et al.*, 2000). A variety of antimicrobial drugs including tetracyclines, penicillins and quinolones have traditionally been used for prophylactic and therapeutic purposes or as growth supplements for food-animal stocks (Angulo *et al.*, 2004; Fabrega *et al.*, 2008). As a consequence, resistance to multiple antimicrobials including fluoroquinolones and expanded-spectrum cephalosporins has emerged among salmonellae that counteract medical treatment of life threatening invasive *Salmonella* infections (Gay *et al.*, 2006; Angulo *et al.*, 2000; Threlfall *et al.*, 2000).

Antimicrobial resistance to a particular drug can be caused by point mutations in the chromosomal DNA and/or by the acquisition of mobile genetic elements such as plasmids, transposons and integrons (Miko *et al.*, 2005). Resistance to nalidixic acid and reduced susceptibility to fluoroquinolones can be mediated by one or more point mutations in the quinolone resistance determining region of the chromosomal *gyrA* gene (Fabrega *et al.*, 2008). Acquisition of resistance to broad-spectrum cephalosporins is associated with plasmid-borne *ampC* genes which encode β-lactamases that hydrolyze the β-lactam ring of these antibiotics (Winokur *et al.*, 2000; Carattoli *et al.*, 2002).

Resistance of *S. enterica* to various antibiotics has increased dramatically in the last two decades (Threlfall *et al.*, 2000). The widespread dissemination of multi-drug resistant (MDR) strains has made it necessary to determine antimicrobial susceptibility patterns in order to take proper prophylactic measures (Threlfall *et al.*, 2000). *Salmonella* isolates are tested against a panel of antibiotics either by the disc diffusion or broth micro-dilution methods, following standard protocols as recommended by the Clinical and Laboratory Standards Institute, U.S.A. (CLSI, formerly known as NCCLS; homepage http://clsi.org/).

In the disc diffusion method, discs containing known amounts of antimicrobials are placed on the agar surface of non-selective media inoculated with the bacterial suspension. A zone appears around the disc after the incubation due to growth inhibition by the antimicrobial that diffuses into the medium from the discs. The width of the inhibition zones are used to classify the isolates as being susceptible, intermediate or resistant to an antimicrobial agent according to the break points as suggested by the CLSI (NCCLS, 2002). For the broth-microdilution method, bacterial suspensions are mixed with non-selective broth and dispensed to the wells of a microtiter plate containing different amounts of various lyophilized antimicrobial agents. After incubation, the plates are screened for bacterial growth, which indicate the minimum inhibitory concentration (MIC in µg/ml) (NCCLS, 2002).

The list of antimicrobials that are tested varies from country to country based on prevalence patterns of individual antimicrobial resistances. The antimicrobials suggested for testing by CLSI include amikacin, ampicillin, amoxicillin-clavulanic acid, cefoxitin, ceftriaxone, ceftiofur, cephalothin, cifrofloxacin, chloramphenicol, gentamicin, kanamycin, nalidixic acid, streptomycin, sulfamethoxazole, tetracycline, and trimethoprim-sulfamethoxazole (NCCLS, 2002).

Antimicrobial susceptibility typing has been useful for *Salmonella* surveillance studies. It has been widely used to study the emergence and spread of antimicrobial resistance in

various *Salmonella* serovars (Threlfall *et al.*, 2003; Thorsteinsdottir *et al.*, 2007). However, antimicrobial sensitivity reflects phenotypic expression and does not allow inferring the genetic descent of isolates because similar phenotypic patterns do not necessarily indicate a common genetic ancestry (Achtman, 2002). Resistance to a particular antimicrobial by an isolate can result by the acquisition of a mobile element from another isolate of a different serovar, e.g., resistance to extended spectrum cephalosporins is plasmid borne and can be horizontally disseminated between different serovars by conjugation (Dunne *et al.*, 2000; Carattoli *et al.*, 2002).

1.6.2 Plasmid profiling

Plasmids are extra-chromosomal DNA molecules that can encode extra-chromosomal virulence and/or antimicrobial resistance (Baker *et al.*, 2007; Hopkins *et al.*, 2008; Carattoli *et al.*, 2002). Plasmid profiling was introduced to investigate the epidemiology of nosocomial Gram-negative bacilli (Schaberg *et al.*, 1981). Bacterial isolates are examined for the presence, number and size of plasmids harboured by them as indicated by the patterns of plasmid DNAs after size separation through electrophoresis on an agarose gel. Alternatively, plasmids can be digested with restriction enzymes before separation for additional resolution (Nauerby *et al.*, 2000).

The simplicity of agarose gel electrophoresis and the speed of this method have made it one of the favourite methods for studying short term epidemiology. Plasmid profiling has been used for intra-serovar differentiation and the source identification of *Salmonella* outbreaks (Taylor *et al.*, 1982; Threlfall *et al.*, 1989; Threlfall *et al.*, 1992). However, not all *Salmonella* isolates carry plasmids and plasmid free strains can not be typed by this method (Foley *et al.*, 2007). Furthermore, plasmids are unstable genetic markers that can be gained or lost. Therefore, plasmid profiling is not suitable for evolutionary studies. Different degrees of supercoiling which produce multiple bands can confound the analysis (Foley *et al.*, 2007). And although restriction based plasmid profiling can resolve among isolates with one or similar numbers of plasmids, the lack of the restriction sites for particular enzymes in some plasmids may render them untypeable.

1.6.3 Ribotyping

For ribotyping, genomic DNA is digested with restriction enzymes, separated by gel electrophoresis and hybridized to DNA probes for rRNA genes (Wachsmuth *et al.*, 1991). Although rRNA genes are highly conserved, they are flanked by variable regions which allows discrimination between strains (Bouchet *et al.*, 2008). This approach has been used in many epidemiological studies of various subspecies *enterica* serovars (Esteban *et al.*, 1993; Liebana, 2002). The method is highly reproducible and can be automated (Foley *et al.*, 2007). Multiple restriction enzymes can be used to increase the discriminatory ability of the method.

The method has limited discrimination for some serovars due to a limited number of rRNA genes (Foley *et al.*, 2007).

1.6.4 IS*200* typing

IS*200* was identified as a *Salmonella* specific insertion element that is approximately 700 bp in size (Lam and Roth, 1983). IS*200* typing of *Salmonella* has been used since 1991 to identify evolutionary lineages within phage types of various subspecies *enterica* serovars (Stanley *et al.*, 1991; Liebana, 2002). IS*200* typing is based on the copy number and location of this insertion element in the *Salmonella* genome. Genomic DNA is digested with one restriction enzyme and the copy number of this insertion element is estimated after southern blotting and hybridization to radiolabelled probes (Helmuth and Schroeter, 1994).

Laborious hybridization has been replaced by IS*200*-PCR, in which IS*200* insertions are amplified using outward-facing primers complementary to the ends of IS*200* (Millemann *et al.*, 2000). The amplified PCR products are separated by agarose gel electrophoresis and the patterns are defined based on the size and number of bands. The method is simpler, rapid and cheaper than restriction-hybridization based IS*200* typing. IS*200* typing has been useful for epidemiological studies (Liebana, 2002). However, this method can not be used for serovars that lack IS*200* element, e.g. Hadar (Millemann *et al.*, 2000).

Restriction–hybridization based methods like IS*200* typing and ribotyping depend on the restriction enzymes that are used as well as the gene probes. Unfortunately, there is a current lack of consensus between different laboratories on which restriction enzymes and probes should be used (Threlfall and Peters, 2001; Liebana, 2002). Furthermore, the discriminatory power of these methods varies with serovar, preventing their use as a general tool for all *Salmonella* serovars.

1.6.5 Pulse field gel electrophoresis (PFGE)

PFGE was developed to separate DNA molecules of 30-2000 Kb on agarose gels for molecular karyotyping of *Saccharomyces cerevisiae* (Schwartz and Cantor, 1984). In PFGE, lysis of bacterial cells and restriction digestion of genomic DNA are carried out within agarose plugs in order to prevent shearing and the fragments are separated by counter-clamped homogenous electric field gel electrophoresis (Centres for Disease Control and Prevention, 2007). This method was introduced to analyze both sporadic and outbreak isolates of subspecies *enterica* serovar Typhi (Thong *et al.*, 1994). PFGE has since been extensively used for the detection and investigation of *Salmonella* outbreaks. Due to the high discrimination of strains in comparison to traditional methods like phage typing and other restriction based methods, PFGE had become the gold standard for epidemiological studies of *Salmonella* by 2001 (Threlfall and Peters, 2001). A molecular database of PFGE profiles of various *Salmonella* serovars is maintained by Centres for Diseases Control and Prevention (CDC), USA that is accessible to epidemiologists and scientists who are the part of the PFGE based network PulseNet (Gerner-Smidt *et al.*, 2006). PulseNet has established a common protocol for the rapid production of standardized and electronically readable fragment patterns.

PFGE is labour and time intensive and needs expertise in performing the procedure and interpreting the results (Foley *et al.*, 2007). Some strains may appear identical following analysis with one restriction enzyme; therefore, multiple PFGE runs are required using different restriction enzymes for better resolution (Nauerby *et al.*, 2000).

1.6.6 Random amplified polymorphic DNA (RAPD)

RAPD is a PCR based method where arbitrary sets of random primers that can bind to the genome at multiple sites are used for the amplification. Amplicons of variable size are amplified based on the proximity of pairs of such sites and the fragments are separated by agarose gel electrophoresis (Hilton *et al.*, 1997; Shangkuan and Lin, 1998). RAPD is easy to use, low cost and provides rapid output. The method has been used for strain discrimination in various subspecies *enterica* serovars, including Typhi (Shangkuan and Lin, 1998; Liebana, 2002; Yan *et al.*, 2003).

Low stringency conditions are used for PCR amplification in RAPD and hence, minor quantitative differences in PCR components can noticeably alter the banding patterns introducing batch to batch variation (Foley *et al.*, 2007). Therefore, the major problem with

this method have been the identification of artefacts and a lack of reproducibility (Liebana, 2002).

1.6.7 Amplified fragment length polymorphism (AFLP)

AFLP was introduced as a high resolution combination of restriction digestion and PCR amplification (Vos *et al.*, 1995). Genomic DNA is digested by a set of restriction enzymes, usually including both a frequent cutter and a rare cutter. The fragments are ligated to DNA linkers in order to allow selective amplification with common primers. After amplification, the fragments are resolved by high resolution gel electrophoresis. Florescent labelled primers can be used for selective amplification and the fragments can be separated by automated capillary electrophoresis (Foley *et al.*, 2007). The technique offers high reproducibility and resolution and no prior knowledge of DNA sequences is required for amplification (Vos *et al.*, 1995; Foley *et al.*, 2007). AFLP has been used as a high resolution fingerprinting method for epidemiological studies of *Salmonella* (Aarts *et al.*, 1998; Gebreyes and Altier, 2002; Torpdahl and Ahrens, 2004; Torpdahl *et al.*, 2005; Gebreyes *et al.*, 2006; Romani *et al.*, 2007).

Although AFLP provides high level of resolution for inter- and intra-serovar differentiation, genetic changes due to nucleotide substitutions will be undetected unless they occur at a restriction site (Foley *et al.*, 2007). The sensitivity of the method also depends on the set of restriction enzyme used and the lack of a common protocol often leads to variable results.

1.6.8 Multi-locus variable number of tandem repeats analysis (MLVA)

Variable number of tandem repeats (VNTR) are a class of repetitive DNA that are variable in length and numbers between individual strains (Van Belkum, 1999). These elements are both ubiquitous and heterogeneous among prokaryotic species (Van Belkum *et al.*, 1998; Van Belkum, 1999), making them a favourite target to study inter-strain polymorphism within a species. The availability of genome sequences and computer programs that can identify VNTRs have made this method very popular (Ramisse *et al.*, 2004). VNTRs can be amplified in a multiplex PCR using florescent labelled primers and can be separated by capillary electrophoresis (Lindstedt *et al.*, 2003; Lindstedt *et al.*, 2004). This approach has been used for intra-serovar strain differentiation for various *Salmonella* serovars (Liu *et al.*, 2003; Lindstedt *et al.*, 2003; Lindstedt *et al.*, 2004; Ramisse *et al.*, 2004; Cho *et al.*, 2007; Hopkins *et al.*, 2007). MLVA typing offers reproducibility and high discriminatory power

that is comparable to PFGE. The technique is easy to perform and analysis of the results is relatively simple (Lindstedt *et al.*, 2004).

MLVA is particularly useful for outbreak investigations and can distinguish between the strains of distinct outbreaks that are associated with a single phage and/or PFGE type (Hopkins *et al.*, 2007). Although VNTR profiles were found to be stable during an outbreak, small changes have been observed due to insertion or deletion of repeats at various loci (Hopkins *et al.*, 2007). VNTRs can evolve very rapidly, resulting in the emergence of multiple profiles during an outbreak that can sometimes lead to erroneous conclusions.

1.6.9 DNA-microarrays for *Salmonella* typing

Whole genome hybridization to multiple probes has been used to compare the gene content among *Salmonella* isolates at both inter- and intra-serovar levels (Porwollik *et al.*, 2002; Boyd *et al.*, 2003; Chan *et al.*, 2003; Porwollik and McClelland, 2003; McClelland *et al.*, 2004; Porwollik *et al.*, 2004). In subsequently recent analyses, the technique has been based on oligonucleotide probes of 20-40 bps that can detect genes encoding antibiotic resistance, pathogenicity, fimbriae, phage-associated genes, flagellae (H-antigens) and lipopolysacharides (O-antigens) (Malorny *et al.*, 2007; Yoshida *et al.*, 2007; Majtan *et al.*, 2007). This technique has been suggested as a powerful tool for strain characterization and epidemiological studies. A high level of accuracy has been validated by gene specific PCRs and other phenotypic methods (Malorny *et al.*, 2007; Majtan *et al.*, 2007). However, uncertainties and discordance in results have also been noticed for a small fraction of data points (Malorny *et al.*, 2007; Yoshida *et al.*, 2007; Majtan *et al.*, 2007).

The approach sounds promising for the reliable characterization of strains in comparison to other existing techniques. However, current analyses have described only small sample sizes from selected serovars. Therefore, the approach still should be validated by testing more diverse sets of larger sample size. Furthermore, the technique is expensive and time and labour intensive.

In summary, serotyping is the first step of strain identification and primary characterization that is followed by other phenotypic methods like phage typing and antimicrobial susceptibility typing. Various genotyping techniques are then employed to further discriminate strains, such as plasmid profiling, ribotyping, IS*200* typing, RAPD, PFGE, AFLP and MLVA. These typing methods are particularly useful for outbreak

characterization and have variable strengths and limitations. A method should have high discriminatory power and reproducibility for strain discrimination and outbreak characterization. Multiple methods are usually required to sufficiently discriminate between the isolates of a serovar (Foley *et al.*, 2007) because strains that are indistinguishable by one method can sometimes be differentiated using others. A typing method should be chosen carefully based on the serovar that has to be analyzed. Although the results of ribotyping are highly reproducible, only few rRNA genes exist per genome, resulting in limited resolution for some serovars (Foley *et al.*, 2007). Methods such as plasmid profiling and IS*200* typing cannot be used for serovars that lack plasmids and insertion element (Millemann *et al.*, 2000; Foley *et al.*, 2007). RAPD offers slightly higher resolution but reproducibility has been a major issue (Liebana, 2002). AFLP, MLVA and PFGE provide the highest resolution but it is very difficult to currently state which one is best. The discriminatory powers of these methods depend on the serovar or phage type (Foley *et al.*, 2007). However, the powers of these methods can be improved ad hoc by using additional restriction enzymes in PFGE and including more loci in MLVA analysis (Foley *et al.*, 2007).

1.7 Population genetics approaches

Several methods such as PFGE, AFLP and MLVA have been used to detect microvariation within bacterial populations and outbreak studies (also known as short term or local epidemiology). However, these methods may not be suitable to study the relatedness amongst the strains from different geographic regions (also called global epidemiology) (Maiden *et al.*, 1998). In contrast, techniques that identify neutral variation may be preferable for studying the evolution and population structure of a bacterial population (Maiden *et al.*, 1998), because rapidly evolving variation usually indicate diversifying selection and might not reflect true strain relatedness.

1.7.1 Multilocus enzyme electrophoresis (MLEE)

MLEE was a standard method in eukaryotic population genetics that was adapted to study evolutionary relationship among different bacterial populations. MLEE identifies slowly accumulating variation in metabolic enzymes that is likely to be selectively neutral (Selander *et al.*, 1986). In MLEE, metabolic enzymes are subjected to electrophoretic separation followed by selective enzyme staining (Selander *et al.*, 1986). The mobilities of

enzymes from different isolates are compared and electromorphs (allozymes or mobility variants) of each enzyme are assigned allele numbers. An absence of enzyme activity is scored as a null allele. Each distinctive combination of alleles is assigned a numeric electrophoretic type (ET).

MLEE depends on differences in electrophoretic mobility that depend on the net electrostatic charge which can change with the amino acid sequence. Only changes in charged amino acids alter electrophoretic mobility. The data is analyzed by phylogenetic and statistical methods to study genetic relatedness among the isolates of a species. MLEE was used to study evolution and population structure of various *Salmonella* serovars in 1980s and 1990s predominantly by Selander's group (Beltran *et al.*, 1988; Boyd *et al.*, 1996; Selander *et al.*, 1992). In 1988, Beltran *et al*. analyzed isolates of eight most common serovars of subspecies *enterica* and observed strong linkage disequilibrium among the enzyme loci, indicating a clonal population structure for the subspecies (Beltran *et al.*, 1988). Single predominant clones were observed for Choleraesuis, Dublin, Heidelberg and Typhimurium whereas multiple evolutionary lineages were reported for serovars Derby, Enteritidis, Infantis, and Newport. Isolates of serovar Typhi were genetically indistinguishable to each other but distinct from other *enterica* serovars (Reeves *et al.*, 1989; Selander *et al.*, 1990b). Three lineages were observed for serovar Paratyphi B, one of which contained isolates of diverse biochemical subtypes and was globally distributed (Selander *et al.*, 1990a). Isolates of serovars Paratyphi A and Sendai were genetically closely related. Serovars Paratyphi C and Panama were monophyletic, i.e. all isolates were members of a single predominant clone for each of these serovars (Selander *et al.*, 1990b). Multiple evolutionary lineages were observed within serovars Typhisuis and Decatur. Most isolates of serovar Dublin were closely related to serovar Enteritidis (Selander *et al.*, 1992). However, five atypical isolates of serovar Dublin were more close to isolates of Paratyphi B and Agona than other Dublin isolates.

Each of the *S. enterica* subspecies was found to represent a distinct group of isolates (Reeves *et al.*, 1989). *S. bongori* that was formerly considered to be a subspecies of *S. enterica* represented a distinct biological species (Reeves *et al.*, 1989). *S. enterica* was proposed to have a clonal population structure due to strong linkage disequilibrium between the loci. And the acquisition of genes encoding antigenic determinants by lateral gene transfer and recombination was suggested to be the source of occurrence of multiple lineages within a serovar (Beltran *et al.*, 1988).

Three *Salmonella* reference collections (SARA, SARB and SARC) were set up based on strain characterization by MLEE. SARA, a collection of 72 isolates of serovars Typhimurium, Saintpaul, Heidelberg, Paratyphi B and Muenchen, was set up in 1991 that represented the full range of genotypic variation in these serovars at that time (Beltran *et al.*, 1991). Another collection, SARB, represented the genetic and phenotypic variation within all of subspecies *enterica.* A total of 72 isolates of 37 subspecies *enterica* serovars were chosen to represent 71 distinct ETs (Boyd *et al.*, 1993). Sixteen *Salmonella* isolates that were characterized by MLEE and nucleotide sequencing of five housekeeping and seven invasion genes were designated as SARC (Boyd *et al.*, 1996). This collection included two representatives from *S. bongori* and each of the *S. enterica* subspecies with two additional isolates from a biotype of subspecies *diarizonae.*

MLEE can distinguish between the subspecies of *S. enterica* as well as the serovars within subspecies *enterica*. However, the resolution may not be sufficient for short-term epidemiology. PFGE has much higher discriminatory powers than MLEE to resolve between the strains of a highly uniform serovar like Typhi and would be more appropriate for outbreak investigations (Thong *et al.*, 1994). MLEE is based on the loci that are likely to be selectively neural and is more suitable for evolutionary studies. However, major disadvantages that are associated with MLEE include a lack of reproducibility between different laboratories, the need to include standards for all known variants of each enzyme locus in every analysis and a lack of globally accessible web-based databases. Finally, MLEE cannot detect amino acid polymorphism that does not result in changed charge or synonymous sequence changes which are the most frequent source of genetic polymorphism.

1.7.2 Multilocus sequence typing (MLST)

MLST was developed in 1998 to apply the proven concept of MLEE to DNA sequences, i.e. the detection of slowly accumulated sequence variation at selectively neutral loci (Maiden *et al.*, 1998). The alleles are directly identified from the nucleotide sequence of the fragments (400-600 bp) of housekeeping genes rather than by comparing the electrophoretic mobility of metabolic enzymes (Maiden *et al.*, 1998). In an MLST scheme, housekeeping genes are selected that are scattered around the bacterial genome. One fragment of each housekeeping gene is amplified and sequenced. Every unique sequence is given a distinct allele number and each unique combination of alleles is assigned a distinct sequence type (ST) number (Maiden *et al.*, 1998).

Many drawbacks of MLEE were overcome by MLST. The results are highly reproducible and the data are archived in the form of the Web based databases in order to enable easy exchange of the information. MLST can be automated for high-throughput. The alleles identified by MLST would be far more in numbers in comparison to MLEE because only five percent of total genetic changes alter electrophoretic properties of metabolic enzymes (Maiden, 2006). The availability of complete genome sequences has been useful for designing MLST schemes with rationally chosen genes. The reduction of sequencing costs has also facilitated this approach. The approach has extensively been used for a wide range of pathogenic and non-pathogenic organisms and led to an improved understanding of the evolution and population biology of these organisms (Maiden, 2006). The wide acceptance of this approach has resulted in more than 48 MLST schemes for various organisms (http://web.mpiib-berlin.mpg.de/mlst/, http://pubmlst.org/ and http://www.mlst.net/).

Three MLST schemes have been proposed to study the evolutionary relatedness among salmonellae. Kotetishvili *et al.* proposed an MLST scheme in 2002 for *Salmonella* typing that was based on four genes: 16S RNA, phosphomannomutase (*manB*), glutamine synthetase (*glnA*) and the 1,2-propanediol utilization factor (*pduF*) (Kotetishvili *et al.*, 2002). The scheme was designed for the epidemiological investigation of *Salmonella* outbreaks and was found to be more discriminative than PFGE (Kotetishvili *et al.*, 2002). Both clinical and environmental *Salmonella* isolates were analyzed by this scheme and various genes were found to be evolving at different rates via different evolutionary mechanisms, including recombination (Kotetishvili *et al.*, 2002). However, typing only four genes may limit the future use of this scheme for large numbers of isolates. After testing twelve loci for meningococcus, seven were chosen to provide sufficient resolution and reliable conclusions (Maiden, 2006). Most of MLST schemes are based on 6 to 10 loci (Maiden, 2006).

Another MLST scheme involved the sequencing of five housekeeping namely *panB* (ketopentoate hydroxymethyltransferase), *icd* (isocitrate dehydrogenase), *manB* (phosphomannomutase), *mdh* (malate dehydrogenase) and *aceK* (isocitrate dehydrogenase kinase/phsophatase) and two virulence genes *fimA* (fimbrial gene A) and *spaN* (surface antigen) (Sukhnanand *et al.*, 2005). Serovar specific STs were observed for various subspecies *enterica* isolates. However, this scheme may have limited implications in evolutionary studies of *Salmonella* because it includes two virulence genes that may be under positive selection, making it more suitable for outbreak characterization. The authors

also proposed an even more limited three gene subtyping scheme based on *manB*, *fimA* and *mdh* for serovar prediction and limited subtype discrimination.

To study the evolution and population structure of serovar Typhi, a global collection of isolates was analyzed by sequencing fragments of seven housekeeping genes, *aroC* (chorismate synthase), *dnaN* (DNA polymerase III beta subunit), *hemD* (uroporphyrinogen III cosynthase), *hisD* (histidinol dehydrogenase), *purE* (phosporibosylaminoimidazole carboxylase), *sucA* (α-ketoglutarate dehydrogense) and *thrA* (aspartokinase and homoserine dehydrogenase) (Kidgell *et al.*, 2002). All the Typhi isolates were genetically uniform, suggesting that they descended from a common ancestor 15,000-150,000 years ago (Kidgell *et al.*, 2002). The genome of Typhi strain CT18 was used for selecting the seven housekeeping genes that are scattered around the genome and they encode the metabolic enzymes that are unlikely to be under diversifying selection (Kidgell *et al.*, 2002).

The scheme fits the requirement of a standard MLST scheme and was later adapted to all *Salmonella* (Torpdahl *et al.*, 2005). Subspecies *enterica* isolates from humans and veterinary sources were analyzed by this scheme and the results were compared to those of PFGE and AFLP. Although a consistency was observed among the genetic trees on the data generated by the three methods, PFGE and AFLP were found to be more discriminatory (Torpdahl *et al.*, 2005). Similar conclusion was reached when isolates of serovar Newport were analyzed by this MLST scheme, PFGE and antimicrobial susceptibility typing, indicating that this scheme may have limited implications for epidemiological investigations (Harbottle *et al.*, 2006). MLST data has also been used to infer the evolution and speciation in *Salmonella* (Falush *et al.*, 2006). Recombination was found to be significant within subspecies *enterica*. However, genetic exchanges between subspecies *enterica* and other subspecies as well as *S. bongori* has been rare (Falush *et al.*, 2006). Although sequential splitting of all the subspecies from a common ancestor was not completely discerned, they seem to have diverged by multiple speciation events and correspond to genetically distinct biological groups (Falush *et al.*, 2006). This scheme has been widely accepted by many members of the *Salmonella* community (Torpdahl *et al.*, 2005; Harbottle *et al.*, 2006; Falush *et al.*, 2006), resulting in > 2350 *Salmonella* strains and > 500 STs on the *Salmonella* MLST website (http://web.mpiib-berlin.mpg.de/mlst/dbs/S*enterica*).

1.8 Population structure

In a bacterial population, clones are defined as groups of genetically indistinguishable isolates that are asexually descended from a common ancestor (Spratt, 2004). Most of the genetic variation in these groups is introduced by mutation events rather than recombination, resulting in a low recombination-mutation rate ratio (Hanage *et al.*, 2006; Turner *et al.*, 2007). The population structure of most of bacterial species was thought to be clonal (Achtman *et al.*, 1983; Selander *et al.*, 1987) until 1993 when Maynard-Smith *et al.* showed that they could vary from strictly clonal to highly sexual (Maynard Smith *et al.*, 1993). *S. enterica* was thought to have a clonal population structure as demonstrated by high values of linkage disequilibrium on MLEE data whereas *Neisseria gonorrhoeae* was found to be highly sexual (Maynard Smith *et al.*, 1993). Clonal population structures have also been reported in genetically monomorphic species like *Mycobacterium tuberculosis* (Baker *et al.*, 2004) and subspecies *enterica* serovar Typhi (Roumagnac *et al.*, 2006). These organisms are designated genetically monomorphic because they harbour very low levels of nucleotide diversity. But genetically monomorphic populations should be distinguished from clonal populations because it would be difficult to estimate true recombination rates in a monomorphic population due to the low levels of nucleotide diversity.

A clonal population structure was deduced for *S. enterica* by researchers in 1980s and 1990s based on MLEE where each distinct serovar evolved from a single or a few predominant clones (Selander *et al.*, 1990b; Beltran *et al.*, 1988; Selander *et al.*, 1994; Reeves *et al.*, 1989). However, recent studies demonstrated that homologous recombination has been the major source of variation among the housekeeping genes within subspecies *enterica* (Brown *et al.*, 2003; Octavia and Lan, 2006; Falush *et al.*, 2006). The phylogenetic trees based on nucleotide sequences found to be incongruent for some of housekeeping genes (Brown *et al.*, 2003; Octavia and Lan, 2006). A mosaic pattern of ancestry and substantial allele sharing between different clades of the phylogenetic tree of subspecies *enterica* indicated significant recombination within this subspecies (Falush *et al.*, 2006). Homologous recombination led to the similarity of gene content and nucleotide divergence in a quarter of genome of Typhi and Paratyphi A, two subspecies *enterica* serovars (Didelot *et al.*, 2007). Therefore, due to significant levels of detectable recombination, phylogenetic approaches may not be suitable to study evolutionary relatedness within subspecies *enterica*.

Homologous recombination and lateral gene transfer played the key role in the evolution of *Salmonella* genome (Porwollik *et al.*, 2002; Porwollik and McClelland, 2003; Porwollik *et al.*, 2004). Almost a quarter of the Typhimurium genome seems to have acquired by lateral gene transfer after the divergence from *Escherichia coli*. However, most of these acquisitions confer virulence and host specificity (Porwollik and McClelland, 2003). Although recombination found to be significant within subspecies *enterica*, genetic exchange between *enterica* and other subspecies is rare (Falush *et al.*, 2006). The studies based on the nucleotide sequences of six housekeeping genes on SARC collection also demonstrated recombination to be rare between subspecies *enterica* and other subspecies (Nelson *et al.*, 1991; Nelson and Selander, 1992; Boyd *et al.*, 1994; Nelson and Selander, 1994; Wang *et al.*, 1997; Nelson *et al.*, 1997). However, phylogenetic incongruence has been observed for some of the housekeeping genes using more sophisticated analyses (Brown *et al.*, 2002). Recombination and lateral gene transfer is the major source of variation among genes coding for lipopolysaccharide and flagellar antigens in *Salmonella* and is also responsible for the occurrence of multiple distinct lineages within a serovar (Li *et al.*, 1994; Selander *et al.*, 1994; Xiang *et al.*, 1994).

1.9 Research goals

S. enterica subsp. *enterica* serovar Newport (henceforth referred to as Newport) has emerged as one of the major human and animal pathogen in recent years. However, limited efforts have been made to study the population structure of this serovar (Beltran *et al.*, 1988; Sukhnanand *et al.*, 2005; Harbottle *et al.*, 2006). Newport was found to have evolved in two lineages by a MLEE study, one associated to humans and the second to animals (Beltran *et al.*, 1988). Similarly, two distinct populations were identified for Newport by studies based on two different MLST schemes (Sukhnanand *et al.*, 2005). However, these conclusions were based on a small number of isolates. Therefore, I analyzed 384 Newport isolates to study the evolution and population structure of serovar Newport, as the first objective of my thesis.

1-5% of the patients with *Salmonella* infection become chronic carriers (D'Aoust, 1991; Gupta *et al.*, 2006). Asymptomatic *Salmonella* carriers are a potential threat to the healthy community, especially those associated with food handling (Ollinger-Snyder and Matthews, 1996; Gupta *et al.*, 2006). It has not yet been determined whether salmonellae from chronic human carriers are the same as those isolated from non-carrier hosts. For the second objective of my thesis, I analyzed 47 isolates from human carriers and compared the

MLST data with that of 1015 isolates from non-carrier humans and animals of the same serovars.

Host associated subtypes have been identified for various subspecies *enterica* serovars (Alcaine *et al.*, 2006; Cho *et al.*, 2007; Rabsch *et al.*, 2002). Although reptiles are one of the largest sources of *Salmonella*, they have been overlooked because most of these studies focused on isolates from humans and domesticated animals. I analyzed 137 subspecies *enterica* isolates belonging to 23 serovars that were isolated from reptiles for the third objective of my thesis. MLST data was compared with that of isolates from humans and non-human warm blooded animals of the same serovars to study whether the clonal diversity in subspecies *enterica* is associated with host categories.

2. Materials

2.1 Equipments and consumables

The names and addresses of the companies, list of chemicals, enzymes and kits, other consumables and equipments used in this study are summarized in Tables 2.1-2.4.

Table 2.1. Names and addresses of companies

Company	Address
ABgene	Epsom, United Kingdom
AccuMed International Limited	East Grinstead, UK
Alcatel	Annecy, France
Applied Biosystems	Foster City, California, USA
Bender and Hobinag	Zurich, Switzerland
BioRad Laboratories	Richmond, California, USA
Calbiochem	San Diego, California, USA
Carl Roth GmbH	Karlsruhe, Germany
Corning Incorporated	New York, U.S.A
Difco Laboratories	Detroit, Michigan, USA
Eppendorf	Hamburg, Germany
Fermentas	Burlington, Ontario, Canada
Genomed GmbH	Loehne, Germany
GFL	Burgwedel, Germany
Heraeus Instruments	Hanau, Germany
IKA-WERK GmbH & Co. KG	Staufen, Germany
INFROS HT	Bottmingen, Switzerland
Invitrogen	Carlsbad, California, USA
Jouan SA	St. Herblain, France
LEGACI Refrigeration Systems	Asheville, North Carolina, USA
Merck	Darmstadt, Germany
Mettler Waagen GmbH	Gissen, Germany
Molecular Devices	Sunnyvale, California, USA
MPIMG	Berlin, Germany
Nunc	Roskilde, Denmark
Oxoid	Cambridge, UK
Pharmacia Biotech	Uppsala, Sweden
Sartorius AG	Goettingen, Germany
SIGMA-ALDRICH Chemie GmbH	Steinheim, Germany
Trek Diagnostics Systems Limited	West Sussex, UK
USB corporation	Cleveland, Ohio, USA

Table 2.2. List of reagents, kits and enzymes

Chemicals/Kits/enzymes	Catalogue no.	Company
Acetic acid	1.00063.2500	Merck
Antimicrobial discs Amikacin	CT0107B	Oxoid
Antimicrobial discs Cephalothin	CT0010B	Oxoid
Antimicrobial discs Cefoxitin	CT0119B	Oxoid
Antimicrobial discs Ceftriaxone	CT0417B	Oxoid
Agarose	R0491	Fermentas
Bacto- yeast extract	0127-01-7	Difco Laboratories

Table 2.2 continued…

Kits/Reagents/enzymes	Catalogue no.	Company
Bacto-Tryptone	0123-01	Difco Laboratories
Bacto-Agar	0140-01	Difco Laboratories
Big dye terminator v 3.1 kit	4337958	Applied Biosystems
Bromophenol blue	1.01895.0010	Merck
Di-sodium hydrogem phosphate	1.06580.1000	Merck
DNA extraction Kit (JetFlex)	600500	Genomed GmbH
DNA polymerase		MPIMG
EDTA	1.08418	Merck
Ethidium bromide	331565	Calbiochem
Ethyl alcohol	1.00983.2511	Merck
Exonuclease I	70073X	USB
Hi Di formamide	4311320	Applied Biosystems
Iso-propanol	9866-1	Carl Roth GmbH
Magnesium chloride	5833	Merck
Microtitre plates with antimicrobials	NLMV1A	Trek Diagnostics Systems Ltd.
Molecular weight marker	10787-018	Invitrogen
Mono-potassium phosphate	1.04873.1000	Merck
Müller-Hinton agar	CM0337	Oxoid
Müller-Hinton broth	CM0405	Oxoid
Potassium chloride	4936-1000	Merck
Shrimp Alkaline Phosphatase	70092Y	USB
Skimmed milk	0032-01-1	Difco Laboratories
Sodium acetate	6268	Merck
Sodium chloride	1.06404.5000	Merck
Sodium hydroxide	1.06495.1000	Merck
Sucrose	7653	Merck
Tris	1.08382.1000	Merck
Trizma base	T-1503	SIGMA-ALDRICH Chemie GmbH
Trizma hydrochloride	T-3253	SIGMA-ALDRICH Chemie GmbH

Table 2.3 Other consumables

Other consumables	Catalogue no.	Company
Glass beads (Φ=2.8-3.3 mm)	A557.1	Carl Roth GmbH
Pasteur pipettes	A4522.1	Carl Roth GmbH
Cryo-tubes (1.8 ml)	375418	Nunc GmbH & Co.
96-well propylene microplates	Costar-6551	Corning Incorporation
96-well PCR plate	AB-0600	ABgene
96-deep well plate (2 ml)	Costar-3960	Corning Incorporation
Plate sealing film	AB-0558	ABgene

Table 2.4. Equipments

Equipment	Model	Company
Centrifuge	5810R	Eppendorf
Centrifuge	5415D	Eppendorf
Deep freezer (-80°C)	Ultima II	LEGACI Refrigeration Systems
DNA analyzer	ABI PRISM 3730	Applied Biosystems
Electrophoresis unit (Power supply & gel chambers)		MPIMG workshop
Gel documentation system	Gel Doc 2000	BioRad Laboratories

Table 2.4 continued …

Equipment	Model	Company
Incubation oven	B6120	Heraeus Instruments
Incubation shaker	RS-T	INFROS HT
Laminar airflow	Herasafe Plus HS15	Heraeus Instruments
Laminar airflow	Herasafe HS12	Heraeus Instruments
Magnetic Stirrer	IKAMEG RET-G	IKA-WERK GmbH & Co. KG
Sensititre autoinoculator	INO2	Trek Diagnostics Systems Ltd.
SensiTouch system		AccuMed International Ltd.
Spectrophotometer	SpectraMax 190	Molecular Devices
Spectrophotometer	GeneQuant	Pharmacia Biotech
Thermocycler	MJ Research PTC225	BioRad Laboratories
Vaccume centrifuge	RC10.22	Jouan SA
Vaccume pump	PASCAL 2005SD	ALCATEL
Vortex	K550GE	Bender and Hobinag
Waterbath	1001	GFL
Weighing machine	BP2100S	Sartorius AG
Weighing machine	PE360	Mettler Waagen GmbH

2.2 Bacterial strains

Salmonella isolates analyzed for various research goals are summarized in Tables 2.5-2.8.

2.3 Oligonucleotides

List of the primers that were used for the amplification and sequencing of gene fragments are mentioned in Table 2.9.

2.4 Media and solutions

LB medium

10 g Bacto-tryptone, 5 g yeast extract and 10 g NaCl were dissolved in 900 ml of distilled water. The pH was adjusted to 7.0 using 1 M NaOH. Final volume was adjusted to 1 litre using distilled water. The medium was sterilized by autoclaving.

LB agar was prepared by adding 15 g of bacto-agar to above preparation before adjusting the final volume to 1 litre while for semisolid LB agar only 7 g of bacto-agar was used.

10X PCR buffer

500 mM KCl, 100 mM Tris (pH 8.3 at 25°C), 15 mM $MgCl_2$ and 0.01% gelatine

5X sequencing buffer

400 mM Tris (pH 9.0 at 25°C) and 10 mM $MgCl_2$

Table 2.5. List of Newport isolates

Z No.	DNANo.	StrainID	Subspecies	Serotype	Host	Source	Country	ST	Year	Continent
10934	4277	73-2008	I	Newport	Human	W. Rabsch	Germany	166	1973	Europe
10936	4278	74-25	I	Newport	Human	W. Rabsch	Germany	45	1974	Europe
10938	4279	74-101	I	Newport	Human	W. Rabsch	Germany	45	1974	Europe
10940	4280	74-123	I	Newport	Human	W. Rabsch	Germany	31	1974	Europe
10942	4281	74-141	I	Newport	Human	W. Rabsch	Germany	45	1974	Europe
10944	4282	74-205	I	Newport	Human	W. Rabsch	Germany	118	1974	Europe
10946	4283	74-431	I	Newport	Human	W. Rabsch	Germany	166	1974	Europe
10948	4284	74-435	I	Newport	Human	W. Rabsch	Germany	166	1974	Europe
10950	4285	74-165b	I	Newport	Human	W. Rabsch	Germany	166	1974	Europe
10952	4286	74-187	I	Newport	Human	W. Rabsch	Germany	166	1974	Europe
10954	4287	75-588	I	Newport	Human	W. Rabsch	Germany	118	1975	Europe
10956	4288	75-589	I	Newport	Human	W. Rabsch	Germany	118	1975	Europe
10958	4289	75-702	I	Newport	Human	W. Rabsch	Germany	45	1975	Europe
10960	4290	75-1231	I	Newport	Human	W. Rabsch	Germany	46	1975	Europe
10962	4291	75-1261	I	Newport	Human	W. Rabsch	Germany	31	1975	Europe
10964	4292	76-19	I	Newport	Human	W. Rabsch	Germany	45	1976	Europe
10966	4293	76-1725	I	Newport	Human	W. Rabsch	Germany	45	1976	Europe
10968	4294	76-1726	I	Newport	Human	W. Rabsch	Germany	46	1976	Europe
10970	4295	76-1758	I	Newport	Human	W. Rabsch	Germany	31	1976	Europe
10972	4296	76-1779	I	Newport	Human	W. Rabsch	Germany	46	1976	Europe
10974	4297	76-1795	I	Newport	Human	W. Rabsch	Germany	46	1976	Europe
10976	4298	76-1821	I	Newport	Human	W. Rabsch	Germany	46	1976	Europe
10978	4299	76-1896	I	Newport	Human	W. Rabsch	Germany	46	1976	Europe
10980	4300	74-445	I	Newport	Human	W. Rabsch	Germany	46	1974	Europe
10982	4301	76-2016	I	Newport	Human	W. Rabsch	Germany	46	1976	Europe
10984	4302	76-2063	I	Newport	Human	W. Rabsch	Germany	46	1976	Europe
10986	4331	99-8496	I	Newport	Human	W. Rabsch	Germany	158	1999	Europe
10988	4304	00 5187	I	Newport	Human	W. Rabsch	Germany	167	2000	Europe
10990	4305	00 6045	I	Newport	Human	W. Rabsch	Germany	156	2000	Europe

Table 2.5 continued…

Z No.	DNANo.	StrainID	Subspecies	Serotype	Host	Source	Country	ST	Year	Continent
10992	4306	00 6775	I	Newport	Human	W. Rabsch	Germany	118	2000	Europe
10994	4307	01-3000	I	Newport	Human	W. Rabsch	Germany	45	2001	Europe
10996	4308	01-6102	I	Newport	Human	W. Rabsch	Germany	46	2001	Europe
10998	4309	02-1789	I	Newport	Chicken	W. Rabsch	Germany	31	2002	Europe
11000	4310	02-5961	I	Newport	Human	W. Rabsch	Germany	163	2002	Europe
11002	4311	02-6778	I	Newport	Human	W. Rabsch	Germany	167	2002	Europe
11004	4312	02-9281	I	Newport	Poultry	W. Rabsch	Germany	166	2002	Europe
11006	4313	02-9460	I	Newport	Human	W. Rabsch	Germany	166	2002	Europe
11008	4332	02-10269	I	Newport	Human	W. Rabsch	Germany	166	2002	Europe
11010	4333	03-558	I	Newport	Meat	W. Rabsch	Germany	166	2003	Europe
11012	4334	03-788	I	Newport	Animalfeed	W. Rabsch	Germany	157	2003	Europe
11014	4335	03-2193	I	Newport	Human	W. Rabsch	Germany	166	2003	Europe
11016	4336	03-4162	I	Newport	Human	W. Rabsch	Germany	118	2003	Europe
11018	4319	03-7424	I	Newport	Food	W. Rabsch	Germany	165	2003	Europe
11020	4320	03-7713	I	Newport	food	W. Rabsch	Germany	46	2003	Europe
11022	4321	03-8520	I	Newport	Human	W. Rabsch	Germany	118	2003	Europe
11024	4322	04-127	I	Newport	Human	W. Rabsch	Germany	45	2004	Europe
11026	4323	04-5845	I	Newport	Fertilizer	W. Rabsch	Germany	166	2004	Europe
11028	4324	04-7663	I	Newport	Human	W. Rabsch	Germany	166	2004	Europe
11030	4325	05-914	I	Newport	Human	W. Rabsch	Germany	31	2005	Europe
11032	4326	05-1812	I	Newport	Human	W. Rabsch	United Kingdom	164	2005	Europe
11306	4207	SARB36	I	Newport	Human	F. Boyd	USA	5		North America
11308	4208	SARB37	I	Newport	Human	F. Boyd	Mexico	31		North America
11310	4209	SARB38	I	Newport	Snake	F. Boyd	USA	46	1987	North America
11484	4489	98-00280	I	Newport	Snake	R. Helmuth	Germany	45	1997	Europe
11486	4490	98-00499	I	Newport	Reptile	R. Helmuth	Germany	45	1998	Europe
11488	4491	98-02541	I	Newport	TigerPython	R. Helmuth	Germany	46	1998	Europe
11490	4492	98-02865	I	Newport	Snake	R. Helmuth	Germany	118	1998	Europe
11492	4493	98-02915	I	Newport	Tortoise	R. Helmuth	Germany	118	1998	Europe

Table 2.5 continued…

Z No.	DNANo.	StrainID	Subspecies	Serotype	Host	Source	Country	ST	Year	Continent
11494	4494	01-00182	I	Newport	CornSnake	R. Helmuth	Germany	118	2000	Europe
11496	4495	01-00350	I	Newport	GarterSnake	R. Helmuth	Germany	118	2001	Europe
11498	4496	01-02029	I	Newport	Turtle	R. Helmuth	Germany	118	2001	Europe
11500	4497	01-03113	I	Newport	Python	R. Helmuth	Germany	46	2001	Europe
11502	4498	02-00184	I	Newport	Blood Python	R. Helmuth	Germany	46	2002	Europe
11504	4499	02-01825	I	Newport	Snake	R. Helmuth	Germany	118	2002	Europe
11506	4500	03-01771-2	I	Newport	Colubrid Snake	R. Helmuth	Germany	118	2003	Europe
11508	4501	04-00451	I	Newport	Snake	R. Helmuth	Germany	46	2004	Europe
11510	4502	04-02663	I	Newport	Reticulated Python	R. Helmuth	Germany	45	2004	Europe
11512	4503	05-00763	I	Newport	Snake	R. Helmuth	Germany	45	2005	Europe
11514	4504	05-00842	I	Newport	Snake	R. Helmuth	Germany	118	2005	Europe
11516	4505	05-01286	I	Newport	Gecko	R. Helmuth	Germany	45	2005	Europe
11518	4506	05-01730	I	Newport	Reptile	R. Helmuth	Germany	184	2005	Europe
11520	4507	05-01897	I	Newport	King Snake	R. Helmuth	Germany	118	2005	Europe
11522	4508	05-01994	I	Newport	Snake	R. Helmuth	Germany	118	2005	Europe
11785	4652	00 7093	I	Newport	Human	F-X. Weill	France	45	2000	Europe
11787	4653	00 7325	I	Newport	Human	F-X. Weill	France	45	2000	Europe
11789	4654	00 3525	I	Newport	Human	F-X. Weill	France	45	2000	Europe
11791	4655	00 3767	I	Newport	Human	F-X. Weill	France	45	2000	Europe
11793	4656	02 7891	I	Newport	Human	F-X. Weill	France	45	2002	Europe
11795	4657	00 3784	I	Newport	Human	F-X. Weill	France	45	2000	Europe
11797	4658	00 4165	I	Newport	Human	F-X. Weill	France	45	2000	Europe
11799	4659	00 5089	I	Newport	Human	F-X. Weill	France	45	2000	Europe
11801	4660	01 2288	I	Newport	Human	F-X. Weill	France	45	2001	Europe
11803	4661	04 9597	I	Newport	Human	F-X. Weill	France	45	2004	Europe
11805	4662	03 8748	I	Newport	Human	F-X. Weill	France	118	2003	Europe
11807	4663	03 3184	I	Newport	Human	F-X. Weill	France	45	2003	Europe
11809	4664	03 3349	I	Newport	Human	F-X. Weill	France	45	2003	Europe
11811	4665	03 5145	I	Newport	Human	F-X. Weill	France	45	2003	Europe

Table 2.5 continued…

Z No.	DNANo.	StrainID	Subspecies	Serotype	Host	Source	Country	ST	Year	Continent
11813	4666	03 6521	I	Newport	Human	F-X. Weill	France	45	2003	Europe
11815	4667	03 7338	I	Newport	Human	F-X. Weill	France	45	2003	Europe
11817	4668	03 9969	I	Newport	Human	F-X. Weill	France	45	2003	Europe
11819	4669	05 2439	I	Newport	Human	F-X. Weill	France	166	2005	Europe
11821	4670	00 6147	I	Newport	Human	F-X. Weill	France	166	2000	Europe
11823	4671	00 4093	I	Newport	Human	F-X. Weill	France	156	2000	Europe
11825	4672	00 8034	I	Newport	Human	F-X. Weill	France	166	2000	Europe
11827	4673	01 9032	I	Newport	Human	F-X. Weill	France	166	2001	Europe
11829	4674	04 1701	I	Newport	Human	F-X. Weill	France	132	2004	Europe
11831	4675	00 0448	I	Newport	Human	F-X. Weill	France	166	2000	Europe
11833	4676	04 8765	I	Newport	Human	F-X. Weill	France	166	2004	Europe
11835	4677	01 2174	I	Newport	Human	F-X. Weill	France	156	2001	Europe
11837	4678	01 5348	I	Newport	Human	F-X. Weill	France	156	2001	Europe
11839	4679	04 3489	I	Newport	Human	F-X. Weill	France	156	2004	Europe
11841	4680	03 5595	I	Newport	Human	F-X. Weill	France	156	2003	Europe
11843	4681	03 3813	I	Newport	Human	F-X. Weill	France	31	2003	Europe
11845	4682	03 6344	I	Newport	Human	F-X. Weill	France	46	2003	Europe
11847	4683	03 7463	I	Newport	Human	F-X. Weill	France	46	2003	Europe
11849	4684	04 1198	I	Newport	Human	F-X. Weill	France	31	2004	Europe
11851	4685	04 1534	I	Newport	Human	F-X. Weill	France	45	2004	Europe
11853	4686	04 2006	I	Newport	Human	F-X. Weill	France	45	2004	Europe
11855	4687	04 2417	I	Newport	Human	F-X. Weill	France	46	2004	Europe
11857	4688	04 2487	I	Newport	Human	F-X. Weill	France	166	2004	Europe
11859	4689	04 3976	I	Newport	Human	F-X. Weill	France	45	2004	Europe
11861	4690	04 4849	I	Newport	Human	F-X. Weill	France	118	2004	Europe
11863	4691	00 2611	I	Newport	Human	F-X. Weill	France	45	2000	Europe
11867	4693	408	I	Newport	Human	F-X. Weill	Madagascar	31	1948	Africa
11869	4694	FW4-51	I	Newport	Human	F-X. Weill	France	118	1951	Europe
11871	4695	FW8-52	I	Newport	Human	F-X. Weill	Madagascar	31	1948	Africa

Table 2.5 continued…

Z No.	DNANo.	StrainID	Subspecies	Serotype	Host	Source	Country	ST	Year	Continent
11873	4696	FW50-3	I	Newport	Human	F-X. Weill	Vietnam	31	1946	Asia
11875	4697	Barcelone/34	I	Newport	Pig	F-X. Weill	Spain	46	1953	Europe
11877	4698	FW23-56	I	Newport	Human	F-X. Weill	Vietnam	46	1956	Asia
11879	4699	FW10-56	I	Newport	Snake	F-X. Weill	Vietnam	46	1952	Asia
11881	4700	FW16-56	I	Newport	Gecko	F-X. Weill	Vietnam	46	1952	Asia
11883	4701	FW18-56	I	Newport	Gecko	F-X. Weill	Vietnam	46	1952	Asia
11885	4702	FW24-56	I	Newport	Human	F-X. Weill	Tunisia	46	1956	Africa
11887	4703	FW13-56	I	Newport	Human	F-X. Weill	Vietnam	46	1952	Asia
11889	4704	404	I	Newport	Pig	F-X. Weill	Madagascar	31	1944	Africa
11891	4705	411	I	Newport	Rat	F-X. Weill	Madagascar	31	1948	Africa
11893	4706	FW6-52	I	Newport	Human	F-X. Weill	Congo	31	1952	Africa
11895	4707	FW8-56	I	Newport	Turtle	F-X. Weill	Vietnam	46	1956	Asia
11897	4708	FW9-57	I	Newport	Human	F-X. Weill	Belgium	46	1957	Europe
11899	4709	FW2-58	I	Newport	Turtle	F-X. Weill	Morocco	211	1958	Africa
11901	4710	FW6-58	I	Newport	Turtle	F-X. Weill	Morocco	46	1958	Africa
11903	4711	FW7-58	I	Newport	Human	F-X. Weill	France	118	1958	Europe
11905	4712	FW12-58	I	Newport	Human	F-X. Weill	Algeria	118	1958	Africa
11907	4713	FW1-59	I	Newport	Human	F-X. Weill	France	46	1959	Europe
11909	4714	FW3-59	I	Newport	Human	F-X. Weill	Ethiopia	31	1959	Africa
11911	4715	FW6-60	I	Newport	Human	F-X. Weill	Tunisia	118	1956	Africa
11913	4716	FW7-61	I	Newport	Chameleon	F-X. Weill	Madagascar	46	1957	Africa
11915	4717	FW18-61	I	Newport	Human	F-X. Weill	France	118	1957	Europe
11917	4718	FW3-64	I	Newport	Food	F-X. Weill	France	46	1964	Europe
11919	4719	FW5-64	I	Newport	Human	F-X. Weill	France	31	1964	Europe
11921	4720	FW9-64	I	Newport	Food	F-X. Weill	France	118	1960	Europe
11923	4721	FW39-64	I	Newport	Horse	F-X. Weill	France	166	1960	Europe
11925	4722	FW8-65	I	Newport	Human	F-X. Weill	France	166	1965	Europe
11927	4723	FW26-65	I	Newport	Food	F-X. Weill	France	118	1961	Europe
11929	4724	FW33-65	I	Newport	Human	F-X. Weill	Romania	166	1965	Europe

Table 2.5 continued…

Z No.	DNANo.	StrainID	Subspecies	Serotype	Host	Source	Country	ST	Year	Continent
11931	4725	FW9-66	I	Newport	Lion	F-X. Weill	Venezuela	45	1962	South America
11933	4726	FW10-66	I	Newport	Chicken	F-X. Weill	Venezuela	45	1962	South America
11935	4727	FW19-66	I	Newport	Human	F-X. Weill	Thailand	46	1966	Asia
11937	4728	FW20-66	I	Newport	Human	F-X. Weill	Greece	31	1962	Europe
11939	4729	FW8-67	I	Newport	Pig	F-X. Weill	Brazil	31	1962	South America
11941	4730	50K	I	Newport	Human	F-X. Weill	France	31	1918	Europe
11943	4731	FW15-65	I	Newport	Human	F-X. Weill	Madagascar	46	1965	Africa
12441	4974	SGSC4157	I	Newport		S. Porwollik		375		
12443	4975	SGSC4910	I	Newport		S. Porwollik		45		
12445	4976	SGSC4911	I	Newport		S. Porwollik		5		
12167	4840	3240	I	Newport	Human carrier	R. Curtiss	Germany	166	1985	Europe

Table 2.6. List of isolates from human carriers

Z No.	DNA No.	Strain ID	Subspecies	Serotype	Host	Source	Country	ST	Year	Continent
12139	4826	3225	I	Virchow	Human carrier	R. Curtiss	Germany	16	1985	Europe
12141	4827	3226	I	Infantis	Human carrier	R. Curtiss	Germany	32	1985	Europe
12143	4828	3228	I	Bovismorbificans	Human carrier	R. Curtiss	Germany	142	1985	Europe
12145	4829	3229	I	Anatum	Human carrier	R. Curtiss	Germany	64	1985	Europe
12147	4830	3230	I	Hadar	Human carrier	R. Curtiss	Germany	33	1985	Europe
12149	4831	3231	I	Infantis	Human carrier	R. Curtiss	Germany	32	1985	Europe
12151	4832	3232	I	Anatum	Human carrier	R. Curtiss	Germany	64	1985	Europe
12153	4833	3233	I	Bovismorbificans	Human carrier	R. Curtiss	Germany	142	1985	Europe
12155	4834	3234	I	Paratyphi B var Java	Human carrier	R. Curtiss	Germany	88	1985	Europe
12161	4837	3237	I	Enteritidis	Human carrier	R. Curtiss	Germany	11	1985	Europe
12163	4838	3238	I	Manhattan	Human carrier	R. Curtiss	Germany	18	1985	Europe
12165	4839	3239	I	Ohio	Human carrier	R. Curtiss	Germany	329	1985	Europe
12169	4841	3241	I	Manhattan	Human carrier	R. Curtiss	Germany	18	1985	Europe
12171	4842	3242	I	Heidelberg	Human carrier	R. Curtiss	Germany	15	1985	Europe
12173	4843	3187	I	Typhimurium	Human carrier	R. Curtiss	Germany	19	1985	Europe
12175	4844	3188	I	Typhimurium	Human carrier	R. Curtiss	Germany	19	1985	Europe
12177	4845	3189	I	Typhimurium	Human carrier	R. Curtiss	Germany	19	1985	Europe
12179	4846	3190	I	Typhimurium	Human carrier	R. Curtiss	Germany	19	1985	Europe
12181	4847	3192	I	Typhimurium	Human carrier	R. Curtiss	Germany	19	1985	Europe
12183	4848	3193	I	Typhimurium	Human carrier	R. Curtiss	Germany	19	1985	Europe
12185	4849	3194	I	Typhimurium	Human carrier	R. Curtiss	Germany	19	1985	Europe
12187	4850	3195	I	Typhimurium	Human carrier	R. Curtiss	Germany	19	1985	Europe
12189	4851	3196	I	Typhimurium	Human carrier	R. Curtiss	Germany	19	1985	Europe
12191	4852	3197	I	Typhimurium	Human carrier	R. Curtiss	Germany	19	1985	Europe
12193	4853	3198	I	Typhimurium	Human carrier	R. Curtiss	Germany	19	1985	Europe
12195	4854	3199	I	Typhimurium	Human carrier	R. Curtiss	Germany	19	1985	Europe
12197	4855	3200	I	Typhimurium	Human carrier	R. Curtiss	Germany	19	1985	Europe

Table 2.6 continued…

Z No.	DNA No.	Strain ID	Subspecies	Serotype	Host	Source	Country	ST	Year	Continent
12199	4856	3201	I	Agona	Human carrier	R. Curtiss	Germany	13	1985	Europe
12203	4858	3203	I	Anatum	Human carrier	R. Curtiss	Germany	64	1985	Europe
12205	4859	3204	I	Braenderup	Human carrier	R. Curtiss	Germany	22	1985	Europe
12207	4860	3205	I	Brandenburg	Human carrier	R. Curtiss	Germany	20	1985	Europe
12209	4861	3206	I	Bredeney	Human carrier	R. Curtiss	Germany	241	1985	Europe
12213	4863	3208	I	Derby	Human carrier	R. Curtiss	Germany	39	1985	Europe
12215	4864	3209	I	Hadar	Human carrier	R. Curtiss	Germany	33	1985	Europe
12217	4865	3210	I	Hadar	Human carrier	R. Curtiss	Germany	368	1985	Europe
12221	4867	3212	I	Heidelberg	Human carrier	R. Curtiss	Germany	15	1985	Europe
12223	4868	3213	I	Infantis	Human carrier	R. Curtiss	Germany	32	1985	Europe
12225	4869	3214	I	Infantis	Human carrier	R. Curtiss	Germany	32	1985	Europe
12231	4872	3217	I	Montevideo	Human carrier	R. Curtiss	Germany	138	1985	Europe
12233	4873	3218	I	Ohio	Human carrier	R. Curtiss	Germany	329	1985	Europe
12235	4874	3219	I	Ohio	Human carrier	R. Curtiss	Germany	329	1985	Europe
12237	4875	3220	I	Panama	Human carrier	R. Curtiss	Germany	48	1985	Europe
12239	4876	3221	I	Panama	Human carrier	R. Curtiss	Germany	48	1985	Europe
12241	4877	3222	I	Schwarzengrund	Human carrier	R. Curtiss	Germany	96	1985	Europe
12243	4878	3223	I	Thompson	Human carrier	R. Curtiss	Germany	26	1985	Europe

Table 2.7. List of isolates from reptiles

Z No.	DNA No.	Strain ID	Subspecies	Serotype	Host	Source	Country	ST	Year	Continent
11404	4449	02-00302	I	Braenderup	Snake	R. Helmuth	Germany	22	2001	Europe
11406	4450	03-00124	I	Braenderup	Gecko	R. Helmuth	Germany	22	2002	Europe
11408	4451	03-01323	I	Braenderup	Snake	R. Helmuth	Germany	22	2003	Europe
11410	4452	05-00728	I	Braenderup	Snake	R. Helmuth	Germany	194	2005	Europe
11412	4453	05-02388	I	Braenderup	Monitor Lizard	R. Helmuth	Germany	22	2005	Europe
11414	4454	99-02017-2	I	Dublin	Agame	R. Helmuth	Germany	74	1999	Europe
11416	4455	03-03818	I	Dublin	Snake	R. Helmuth	Germany	180	2003	Europe
11418	4456	99-03574	I	Enteritidis	Snake	R. Helmuth	Germany	11	1999	Europe
11420	4457	00-00610	I	Enteritidis	Snake	R. Helmuth	Germany	11	2000	Europe
11422	4458	01-00493-2	I	Enteritidis	Reptile	R. Helmuth	Germany	168	2001	Europe
11424	4459	02-02908-2	I	Enteritidis	Snake	R. Helmuth	Germany	172	2002	Europe
11426	4460	02-03265	I	Enteritidis	Monitor Lizard	R. Helmuth	Germany	11	2002	Europe
11428	4461	03-01771-1	I	Enteritidis	Colubrid Snake	R. Helmuth	Germany	11	2003	Europe
11430	4462	04-01507	I	Enteritidis	Snake	R. Helmuth	Germany	11	2004	Europe
11432	4463	05-00489	I	Enteritidis	Reptile	R. Helmuth	Germany	180	2005	Europe
11434	4464	98-00123	I	Javiana z28 neg.	Tortoise	R. Helmuth	Germany	175	1997	Europe
11436	4465	02-02216	I	Javiana	Boa Constrictor	R. Helmuth	Germany	24	2002	Europe
11438	4466	02-02877	I	Javiana z28 neg.	Reptile	R. Helmuth	Germany	175	2002	Europe
11440	4467	03-00119	I	Javiana	Snake	R. Helmuth	Germany	24	2002	Europe
11442	4468	02-02256	I	Miami	Reptile	R. Helmuth	Germany	171	2002	Europe
11444	4469	03-01890	I	Miami	Tortoise	R. Helmuth	Germany	171	2003	Europe
11446	4470	98-03021	I	Montevideo	Reptile	R. Helmuth	Germany	4	1998	Europe
11448	4471	98-03023	I	Montevideo	Reptile	R. Helmuth	Germany	4	1998	Europe
11450	4472	98-04046	I	Montevideo	Chameleon	R. Helmuth	Germany	195	1998	Europe
11452	4473	99-00173	I	Montevideo	Monitor Lizard	R. Helmuth	Germany	4	1999	Europe
11454	4474	98-00489	I	Muenchen	Snake	R. Helmuth	Germany	112	1998	Europe
11456	4475	98-00490	I	Muenchen	Snake	R. Helmuth	Germany	112	1998	Europe
11458	4476	98-01002	I	Muenchen	Snake	R. Helmuth	Germany	112	1998	Europe
11460	4477	98-01780	I	Muenchen	Metallic Iguana	R. Helmuth	Germany	112	1998	Europe

Table 2.7 continued…

Z No.	DNA No.	Strain ID	Subspecies	Serotype	Host	Source	Country	ST	Year	Continent
11462	4478	98-04248	I	Muenchen	Tortoise	R. Helmuth	Germany	112	1998	Europe
11464	4479	99-04404	I	Muenchen	Boa Constrictor	R. Helmuth	Germany	112	1999	Europe
11466	4480	01-01300	I	Muenchen	Snake	R. Helmuth	Germany	170	2001	Europe
11468	4481	01-03079	I	Muenchen	Snake	R. Helmuth	Germany	112	2001	Europe
11470	4482	03-00146	I	Muenchen	Snake	R. Helmuth	Germany	112	2002	Europe
11472	4483	03-01321	I	Muenchen	Red Throat Anolis	R. Helmuth	Germany	112	2003	Europe
11474	4484	03-02124	I	Muenchen	Rainbow Boa	R. Helmuth	Germany	112	2003	Europe
11476	4485	04-00724	I	Muenchen	Reptile	R. Helmuth	Germany	176	2004	Europe
11478	4486	04-01589	I	Muenchen	Snake	R. Helmuth	Germany	177	2004	Europe
11480	4487	04-03008-1	I	Muenchen	Snake	R. Helmuth	Germany	178	2004	Europe
11482	4488	04-03596	I	Muenchen	Reptile	R. Helmuth	Germany	173	2004	Europe
11524	4509	98-00501	I	Oranienburg	Reptile	R. Helmuth	Germany	174	1998	Europe
11526	4510	98-01003	I	Oranienburg	Tortoise	R. Helmuth	Germany	23	1998	Europe
11528	4511	98-03150	I	Oranienburg	Lizard	R. Helmuth	Germany	47	1998	Europe
11530	4512	99-01511	I	Oranienburg	Shingleback Lizard	R. Helmuth	Germany	174	1999	Europe
11532	4513	99-03694	I	Oranienburg	Snake	R. Helmuth	Germany	174	1999	Europe
11534	4514	00-02036	I	Oranienburg	Monitor Lizard	R. Helmuth	Germany	23	2000	Europe
11536	4515	00-02043	I	Oranienburg	Malachite Iguana	R. Helmuth	Germany	23	2000	Europe
11538	4516	01-00348	I	Oranienburg	King Python	R. Helmuth	Germany	23	2001	Europe
11540	4517	01-01279	I	Oranienburg	Gecko	R. Helmuth	Germany	23	2001	Europe
11542	4518	01-01632	I	Oranienburg	Reptile	R. Helmuth	Germany	169	2001	Europe
11544	4519	02-00025	I	Oranienburg	Boa Constrictor	R. Helmuth	Germany	179	2002	Europe
11546	4520	02-03361	I	Oranienburg	Bearded Agame	R. Helmuth	Germany	47	2002	Europe
11548	4521	02-04555	I	Oranienburg	Snake	R. Helmuth	Germany	179	2002	Europe
11550	4522	02-04558	I	Oranienburg	Snake	R. Helmuth	Germany	174	2002	Europe
11552	4523	02-04559	I	Oranienburg	Snake	R. Helmuth	Germany	174	2002	Europe
11554	4524	02-04561	I	Oranienburg	Snake	R. Helmuth	Germany	174	2002	Europe
11556	4525	02-04562	I	Oranienburg	Snake	R. Helmuth	Germany	174	2002	Europe
11558	4526	02-04563	I	Oranienburg	Snake	R. Helmuth	Germany	174	2002	Europe

Table 2.7 continued

Z No.	DNA No.	Strain ID	Subspecies	Serotype	Host	Source	Country	ST	Year	Continent
11560	4527	02-04570	I	Oranienburg	Snake	R. Helmuth	Germany	174	2002	Europe
11562	4528	04-00454	I	Oranienburg	Snake	R. Helmuth	Germany	179	2004	Europe
11564	4529	04-01588	I	Oranienburg	Snake	R. Helmuth	Germany	174	2004	Europe
11566	4530	04-03447	I	Oranienburg	Snake	R. Helmuth	Germany	23	2004	Europe
11568	4531	04-03465	I	Oranienburg	Snake	R. Helmuth	Germany	179	2004	Europe
11570	4532	05-00368	I	Oranienburg	Snake	R. Helmuth	Germany	23	2005	Europe
11572	4533	05-00373	I	Oranienburg	Snake	R. Helmuth	Germany	174	2005	Europe
11574	4534	05-00844	I	Oranienburg	Snake	R. Helmuth	Germany	179	2005	Europe
11576	4535	05-01895	I	Oranienburg	Tortoise	R. Helmuth	Germany	23	2005	Europe
11578	4536	05-02633	I	Oranienburg	Snake	R. Helmuth	Germany	179	2005	Europe
11580	4537	05-02834	I	Oranienburg	Snake	R. Helmuth	Germany	179	2005	Europe
11582	4538	05-02969	I	Oranienburg	Snake	R. Helmuth	Germany	174	2005	Europe
11584	4539	99-00287	I	Panama	Snake	R. Helmuth	Germany	48	1999	Europe
11586	4540	99-01236	I	Panama	Snake	R. Helmuth	Germany	48	1999	Europe
11588	4541	02-04627	I	Panama	Corn Snake	R. Helmuth	Germany	48	2002	Europe
11590	4542	04-00740	I	Panama	Snake	R. Helmuth	Germany	48	2004	Europe
11592	4543	04-00939	I	Panama	Snake	R. Helmuth	Germany	48	2004	Europe
11594	4544	00-02650	I	Paratyphi B	Snake	R. Helmuth	Germany	88	2000	Europe
11596	4545	00-03806	I	Paratyphi B	Python	R. Helmuth	Germany	88	2000	Europe
11598	4546	02-00729	I	Paratyphi B	Tortoise	R. Helmuth	Germany	43	2002	Europe
11600	4547	02-04564	I	Paratyphi B	Snake	R. Helmuth	Germany	88	2002	Europe
11602	4548	02-04565	I	Paratyphi B	Snake	R. Helmuth	Germany	88	2002	Europe
11604	4549	02-04626	I	Paratyphi B	Corn Snake	R. Helmuth	Germany	88	2002	Europe
11606	4550	04-01793	I	Paratyphi B	Snake	R. Helmuth	Germany	88	2004	Europe
11608	4551	05-01473	I	Paratyphi B	Snake	R. Helmuth	Germany	88	2005	Europe
11610	4552	02-01364	I	Saintpaul	Tortoise	R. Helmuth	Germany	27	2002	Europe
11612	4553	05-02659	I	Saintpaul	Crocodile	R. Helmuth	Germany	50	2005	Europe
11614	4554	04-00456	I	Senftenberg	Snake	R. Helmuth	Germany	14	2004	Europe
11616	4555	04-03354	I	Senftenberg	Snake	R. Helmuth	Germany	14	2004	Europe

Table 2.7 continued…

Z No.	DNA No.	Strain ID	Subspecies	Serotype	Host	Source	Country	ST	Year	Continent
11618	4556	04-03461	I	Senftenberg	Lizard	R. Helmuth	Germany	14	2004	Europe
11620	4557	01-01318	I	Stanley	Reptile	R. Helmuth	Germany	182	2001	Europe
11622	4558	01-00185	I	Thompson	Turtle	R. Helmuth	Germany	26	2001	Europe
11624	4559	03-02125	I	Thompson	Tortoise	R. Helmuth	Germany	26	2003	Europe
11626	4560	03-02126	I	Thompson	Tortoise	R. Helmuth	Germany	26	2003	Europe
11628	4561	98-02312	I	Typhimurium	Reptile	R. Helmuth	Germany	19	1998	Europe
11630	4562	98-04017	I	Typhimurium	Python	R. Helmuth	Germany	19	1998	Europe
11632	4563	99-03089	I	Typhimurium	Tortoise	R. Helmuth	Germany	19	1999	Europe
11634	4564	01-00493-1	I	Typhimurium	Gecko	R. Helmuth	Germany	34	2001	Europe
11636	4565	04-00449	I	Typhimurium	Snake	R. Helmuth	Germany	19	2004	Europe
11638	4566	99-01235	I	Virchow	Turtle	R. Helmuth	Germany	197	1999	Europe
11640	4567	00-02323	I	Virchow	Four-fingered Tortoise	R. Helmuth	Germany	16	2000	Europe
11642	4568	00-03509	I	Virchow	Lizard	R. Helmuth	Germany	181	2000	Europe
11644	4569	98-00485	I	Anatum	Snake	R. Helmuth	Germany	64	1998	Europe
11646	4570	99-01212	I	Anatum	Reptile	R. Helmuth	Germany	64	1999	Europe
11648	4571	98-01679	I	Infantis	Snake	R. Helmuth	Germany	32	1998	Europe
11650	4572	00-01231	I	Infantis	Snake	R. Helmuth	Germany	32	2000	Europe
11652	4573	98-02313	I	Decatur	Reptile	R. Helmuth	Germany	186	1998	Europe
11654	4574	02-04382	I	Bovismorbificans	Snake	R. Helmuth	Germany	150	2002	Europe
11656	4575	04-00762	I	Bovismorbificans	Reptile	R. Helmuth	Germany	150	2004	Europe
11658	4576	01-01676	I	Brandenburg	Iguana	R. Helmuth	Germany	65	2001	Europe
11664	4579	99-00032	I	Hadar	Snake	R. Helmuth	Germany	33	1999	Europe

Table 2.8. Other *Salmonella* isolates analyzed in this study

Z No.	DNA No.	Strain ID	Subspecies	Serotype	Host	Source	Country	ST	Year	Continent	Phage type
11666	4580	00-03100	I	Enteritidis	Hedgehog	R. Helmuth	Germany	11	2000	Europe	
11668	4581	02-04192	I	Enteritidis	Hedgehog	R. Helmuth	Germany	11	2002	Europe	
11670	4582	04-03542	I	Enteritidis	Hedgehog	R. Helmuth	Germany	183	2004	Europe	
11672	4583	05-02531	I	Enteritidis	Hedgehog	R. Helmuth	Germany	11	2005	Europe	
11674	4584	99-02302	I	Enteritidis	Hedgehog	R. Helmuth	Germany	183	1999	Europe	PT11
11676	4585	00-03508	I	Enteritidis	Hedgehog	R. Helmuth	Germany	183	2000	Europe	PT11
11678	4586	02-00390	I	Enteritidis	Hedgehog	R. Helmuth	Germany	183	2002	Europe	PT11
11680	4587	03-00641	I	Enteritidis	Hedgehog	R. Helmuth	Germany	183	2002	Europe	PT11
11682	4588	04-00150	I	Enteritidis	Hedgehog	R. Helmuth	Germany	183	2003	Europe	PT11
11684	4589	98-00886	I	Senftenberg	Animal Feed	R. Helmuth	Germany	185	1998	Europe	
11686	4590	98-04320	I	Senftenberg	Animal meal	R. Helmuth	Germany	196	1998	Europe	
11688	4591	99-00016	I	Senftenberg	Dog/Cat Food	R. Helmuth	Germany	14	1998	Europe	
11690	4592	99-01195	I	Senftenberg	Cereal	R. Helmuth	Germany	185	1999	Europe	
11692	4593	99-04145-1	I	Senftenberg	Animal Feed	R. Helmuth	Germany	185	1999	Europe	
11694	4594	00-00378	I	Senftenberg	Soya	R. Helmuth	Germany	14	2000	Europe	
11696	4595	00-00822	I	Senftenberg	Animal meal	R. Helmuth	Germany	185	2000	Europe	
11698	4596	00-03768	I	Senftenberg	Animal Feed	R. Helmuth	Germany	192	2000	Europe	
11700	4597	01-00576	I	Senftenberg	Fish meal	R. Helmuth	Germany	14	2001	Europe	
11702	4598	01-02488	I	Senftenberg	Fish meal	R. Helmuth	Germany	14	2001	Europe	
11704	4599	02-00253	I	Senftenberg	Animal Feed	R. Helmuth	Germany	185	2002	Europe	
11706	4600	03-00172	I	Senftenberg	Animal Feed	R. Helmuth	Germany	185	2002	Europe	
11708	4601	03-02853	I	Senftenberg	Dog/Cat Food	R. Helmuth	Germany	14	2003	Europe	
11710	4602	03-02877	I	Senftenberg	Fish Meal	R. Helmuth	Germany	14	2003	Europe	
11712	4603	04-00548	I	Senftenberg	Animal Feed	R. Helmuth	Germany	185	2004	Europe	
11714	4604	04-00997	I	Senftenberg	Soya	R. Helmuth	Germany	14	2004	Europe	
11716	4605	04-01228	I	Senftenberg	Animal Feed	R. Helmuth	Germany	210	2004	Europe	
11718	4606	04-03902	I	Senftenberg	Fish meal	R. Helmuth	Germany	14	2004	Europe	
11720	4607	05-00522	I	Senftenberg	Animal Feed	R. Helmuth	Germany	185	2005	Europe	

Table 2.8 continued…

Z No.	DNA No.	Strain ID	Subspecies	Serotype	Host	Source	Country	ST	Year	Continent	Phage type
11722	4608	05-01371	I	Senftenberg	Animal Feed	R. Helmuth	Germany	14	2005	Europe	
11724	4609	RKI04-07579	I	Kentucky	Human	W. Rabsch	Senegal	198	2004	Africa	
11726	4610	RKI04-07898	I	Kentucky	Human	W. Rabsch	Senegal	198	2004	Africa	
11728	4611	RKI04-07904	I	Kentucky	Human	W. Rabsch	Senegal	198	2004	Africa	
11730	4612	RKI05-02137	I	Kentucky	Falcon	W. Rabsch	United Arab Emirates	198	2005	Asia	
11732	4613	RKI05-02142	I	Kentucky	Cheetah	W. Rabsch	United Arab Emirates	198	2005	Asia	
11734	4614	RKI05-02154	I	Kentucky	Jaguar	W. Rabsch	United Arab Emirates	198	2005	Asia	
11971	4742	R1	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1990	Europe	DT104
11973	4743	R2	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1994	Europe	DT104
11975	4744	R3	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1996	Europe	DT104
11977	4745	R4	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1996	Europe	DT104
11979	4746	R5	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1997	Europe	DT104
11981	4747	R6	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1997	Europe	DT104
11983	4748	R7	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1997	Europe	DT104
11985	4749	R8	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1997	Europe	DT104
11987	4750	R9	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1998	Europe	DT104
11989	4751	R10	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1998	Europe	DT104
11991	4752	R11	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1997	Europe	DT104
11993	4753	R12	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1999	Europe	DT104
11995	4754	R13	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	128	1999	Europe	DT104
11999	4756	R15	I	Typhimurium	Pigeon	H. L A-Polymenis	Germany	128	1992	Europe	DT2
12003	4758	R17	I	Typhimurium	Pigeon	H. L A-Polymenis	Germany	128	1994	Europe	DT2
12005	4759	R18	I	Typhimurium	Pigeon	H. L A-Polymenis	Germany	128	1996	Europe	DT2
12007	4760	R19	I	Typhimurium	Pigeon	H. L A-Polymenis	Germany	128	2000	Europe	DT2
12009	4761	R20	I	Typhimurium	Pigeon	H. L A-Polymenis	Germany	128	2001	Europe	DT2
12011	4762	R21	I	Typhimurium	Human	H. L A-Polymenis	Germany	19	1974	Europe	DT204
12013	4763	R22	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1975	Europe	DT204
12015	4764	R23	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1979	Europe	DT204

Table 2.8 continued…

Z No.	DNA No.	Strain ID	Subspecies	Serotype	Host	Source	Country	ST	Year	Continent	Phage type
12017	4765	R24	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1990	Europe	DT204
12019	4766	R25	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1996	Europe	DT204
12021	4767	R26	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	2000	Europe	DT204
12023	4768	R27	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	2000	Europe	DT204
12025	4769	R28	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1994	Europe	DT204c
12027	4770	R29	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1994	Europe	DT204c
12029	4771	R30	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1997	Europe	DT204c
12031	4772	R31	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1997	Europe	DT204c
12033	4773	R32	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1997	Europe	DT204c
12035	4774	R33	I	Typhimurium	Pigeon	H. L A-Polymenis	Germany	19	1988	Europe	DT99
12037	4775	R34	I	Typhimurium	Pigeon	H. L A-Polymenis	Germany	19	1992	Europe	DT99
12039	4776	R35	I	Typhimurium	Pigeon	H. L A-Polymenis	Germany	19	1994	Europe	DT99
12041	4777	R36	I	Typhimurium	Pigeon	H. L A-Polymenis	Germany	19	1994	Europe	DT99
12043	4778	R37	I	Typhimurium	Pigeon	H. L A-Polymenis	Germany	19	1996	Europe	DT99
12045	4779	R38	I	Typhimurium	Pigeon	H. L A-Polymenis	Germany	19	2000	Europe	DT99
12047	4780	R39	I	Typhimurium	Pigeon	H. L A-Polymenis	Germany	19	2000	Europe	DT99
12049	4781	R40	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1991	Europe	DT49
12051	4782	R41	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1994	Europe	DT49
12053	4783	R42	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1996	Europe	DT49
12055	4784	R43	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1996	Europe	DT49
12057	4785	R44	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1988	Europe	DT193
12059	4786	R45	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1996	Europe	DT193
12061	4787	R46	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1997	Europe	DT193
12063	4788	R47	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1988	Europe	DT15A
12065	4789	R48	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	323	1988	Europe	DT1
12067	4790	R49	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	323	1989	Europe	DT1
12069	4791	R50	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1988	Europe	DT10
12073	4793	R52	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	376	1988	Europe	DT17
12075	4794	R53	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1989	Europe	DT17

Table 2.8 continued…

Z No.	DNA No.	Strain ID	Subspecies	Serotype	Host	Source	Country	ST	Year	Continent	Phage type
12077	4795	R54	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1990	Europe	DT17
12079	4796	R55	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1988	Europe	DT12
12081	4797	R56	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	332	1992	Europe	DT12
12083	4798	R57	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	332	1990	Europe	DT66
12085	4799	R58	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1991	Europe	DT66
12087	4800	R59	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1994	Europe	DT66
12089	4801	R60	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1996	Europe	DT66
12091	4802	R61	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1991	Europe	DT22
12093	4803	R62	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	34	1996	Europe	DT7
12095	4804	R63	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1991	Europe	DT41
12097	4805	R64	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1996	Europe	DT106
12099	4806	R65	I	Typhimurium	Bovine	H. L A-Polymenis	Germany	19	1992	Europe	DT12A
12101	4807	H04	I	Typhimurium	Pigeon	H. L A-Polymenis	Germany	19	1996	Europe	DT99
12103	4808	H05	I	Typhimurium	Pigeon	H. L A-Polymenis	Germany	128	2000	Europe	DT99
12105	4809	H06	I	Typhimurium	Pigeon	H. L A-Polymenis	Germany	19	2000	Europe	DT99
12107	4810	H07	I	Typhimurium	Pigeon	H. L A-Polymenis	Germany	19	2001	Europe	DT99
12109	4811	H08	I	Typhimurium	Pigeon	H. L A-Polymenis	Germany	128	1999	Europe	DT2
12111	4812	H09	I	Typhimurium	Pigeon	H. L A-Polymenis	Germany	128	2000	Europe	DT2
12113	4813	H10	I	Typhimurium	Pigeon	H. L A-Polymenis	Germany	128	2001	Europe	DT2
12115	4814	STM213/94	I	Typhimurium	Pigeon	W. Rabsch	Germany	128	1994	Europe	DT2
12117	4815	STM2350/94	I	Typhimurium	Bovine	W. Rabsch	Germany	19	1994	Europe	DT4
12119	4816	STM3246/94	I	Typhimurium	Bovine	W. Rabsch	Germany	98	1994	Europe	DT36
12121	4817	STM6669/96	I	Typhimurium	Bovine	W. Rabsch	Germany	19	1996	Europe	DT68
12123	4818	STM3244/94	I	Typhimurium	Bovine	W. Rabsch	Germany	19	1994	Europe	DT120
12125	4819	STM499/94	I	Typhimurium	Animal feed	W. Rabsch	Germany	19	1994	Europe	DT120
12127	4820	STM1663/89	I	Typhimurium	Bovine	W. Rabsch	Germany	19	1989	Europe	DT120
12129	4821	STM3115/94	I	Typhimurium	Bovine	W. Rabsch	Germany	19	1994	Europe	DT143
12131	4822	STM1371/94	I	Typhimurium	Bovine	W. Rabsch	Germany	19	1994	Europe	DT170
12133	4823	STM2204/89	I	Typhimurium	Bovine	W. Rabsch	Germany	19	1989	Europe	

Table 2.8 continued…

Z No.	DNA No.	Strain ID	Subspecies	Serotype	Host	Source	Country	ST	Year	Continent	Phage type
12135	4824	STM261/91	I	Typhimurium	Bovine	W. Rabsch	Germany	19	1991	Europe	
12257	4883	MZ0876	I	Gallinarum		S. Porwollik		331	None		
12259	4884	CDC1983-167	I	Pullorum		S. Porwollik		92			
12261	4885	SGSC4770SC-B67	I	Choleraesuis	Human	S. Porwollik	Taiwan	66		Asia	
12265	4886	MZ0526	I	Hadar		S. Porwollik		330	None		
12267	4887	S-1241	I	Paratyphi B var Java		S. Porwollik		372			
12277	4892	ccc28	I	Agona	Milk	S. Porwollik	Ireland	13	2002	Europe	
12291	4899	ATCC 700720	I	Typhimurium		S. Porwollik		19	1948		LT2
12293	4900	MZ0523	I	Javiana		S. Porwollik		371			
12295	4901	S-407	I	Oranienburg		S. Porwollik		320			
12297	4902	MZ0530	I	Oranienburg		S. Porwollik		320	None		
12301	4904	MZ0552	I	Cubana		S. Porwollik		324			
12303	4905	125109	I	Enteritidis		S. Porwollik		11			PT4
12313	4910	SS44	I	Abortusovis	Sheep	S. Porwollik	Italy	202	1980	Europe	
12315	4911	MZ0680	I	Abortusovis		S. Porwollik	France	373	1980	Europe	
12327	4917	SGSC3820	I	Enteritidis	Chicken	S. Porwollik		11			
12403	4955	74-1035	I	Sendai	Human	S. Porwollik	U.S.A.	85		North America	
12407	4957	RKS4594	I	Paratyphi C		S. Porwollik		114			
12409	4958	MZ0955	I	Paratyphi B		S. Porwollik		325	None		
12411	4959	01-05481	I	Bovismorbificans		S. Porwollik	Germany	142		Europe	PT13
12413	4960	05-0774	I	Bovismorbificans		S. Porwollik	Germany	142		Europe	PT24
12415	4961	M1166	I	Typhimurium		S. Porwollik		19			
12417	4962	L14	I	Typhimurium		S. Porwollik		19			
12419	4963	A36	I	Typhimurium		S. Porwollik	Germany	19		Europe	
12421	4964	MZ1292	I	Muenster	Human	S. Porwollik		321	None		
12423	4965	MZ1293	I	Virchow	Human	S. Porwollik		326	None		
12425	4966	04-0258	I	Muenster	Human	S. Porwollik		374			
12427	4967	MZ1295	I	Virchow	Human	S. Porwollik		333	None		

Table 2.8 continued…

Z No.	DNA No.	Strain ID	Subspecies	Serotype	Host	Source	Country	ST	Year	Continent	Phage type
12429	4968	SL1344	I	Typhimurium		S. Porwollik	United Kingdom	19		Europe	
12433	4970	SGSC4902	I	Paratyphi A		S. Porwollik		85			
12435	4971	SGSC4903	I	Typhimurium		S. Porwollik		19			
12437	4972	SGSC4905	I	Infantis		S. Porwollik		32			
12439	4973	MZ1302	I	Hadar		S. Porwollik		327	None		
12447	4977	SGSC4912	I	Schwarzengrund		S. Porwollik		322			
12449	4978	SGSC4913	I			S. Porwollik		19			
12451	4979	SGSC4914	I	Kentucky		S. Porwollik		152			
12453	4980	SGSC4915	I	Heidelberg		S. Porwollik		15			
12455	4981	SGSC4916	I	Dublin		S. Porwollik		10			
12457	4982	SGSC4917	I	Javiana		S. Porwollik		24			
12459	4983	SGSC4918	I	Kentucky		S. Porwollik		152			
12461	4984	MZ1313	I	Schwarzengrund		S. Porwollik		322	None		
12463	4985	SGSC4920	I	Saintpaul		S. Porwollik		50			
12465	4986	SGSC4921	I	Saintpaul		S. Porwollik		95			
12157	4835	3235	I	Rough (4,12:d:-)	Human carrier	R. Curtiss	Germany	279	1985	Europe	
12159	4836	3236	I	Tennessee	Human carrier	R. Curtiss	Germany	319	1985	Europe	
12201	4857	3202	I	Albany	Human carrier	R. Curtiss	Germany	292	1985	Europe	
12211	4862	3207	I	Cerro var. Siegburg	Human carrier	R. Curtiss	Germany	367	1985	Europe	
12219	4866	3211	I	Give	Human carrier	R. Curtiss	Germany	369	1985	Europe	
12227	4870	3215	I	London	Human carrier	R. Curtiss	Germany	155	1985	Europe	
12229	4871	3216	I	London	Human carrier	R. Curtiss	Germany	155	1985	Europe	
12137	4825	3224	I	Vejle	Human carrier	R. Curtiss	Germany	370	1985	Europe	

Note: These isolates were analyzed and MLST data were submitted to the *Salmonella* MLST website (http://web.mpiib-berlin.mpg.de/mlst/dbs/Senterica). The data of relevant isolates were used in comparative analyses for different objectives of this study.

Table 2.9. Primers for the amplification and sequencing of MLST gene fragments

O- number	Gene name	Primer direction	Oligonucleotide sequence
Amplification			
2181	*aroC*	Forward	CCTGGCACCTCGCGCTATAC
2182	*aroC*	Reverse	CCACACACGGATCGTGGCG
2173	*dnaN*	Forward	ATGAAATTTACCGTTGAACGTGA
6586	*dnaN*	Reverse	AATTTCTCATTCGAGAGGATTGC
2169	*hemD*	Forward	ATGAGTATTCTGATCACCCG
6584	*hemD*	Reverse	TTATTGTAATGCGCGCAACAG
6080	*hisD*	Forward	GAAACGTTCCATTCCGCGCAGAC
6081	*hisD*	Reverse	CTGAACGGTCATCCGTTTCTG
6589	*purE*	Forward	ATGTCTTCCCGCAATAATCC
2178	*purE*	Reverse	TCATAGCGTCCCCCGCGGATC
6593	*sucA*	Forward	CCTTTGCCAGCGGCAAAGAGAC
2190	*sucA*	Reverse	CGCATTGACGTGGAAAATCGG
6078	*thrA*	Forward	GTCACGGTGATCGATCCGGT
6079	*thrA*	Reverse	CACGATATTGATATTAGCCCG
Sequencing			
2184	*aroC*	Reverse	CATATGCGCCACAATGTGTTG
6587	*dnaN*	Forward	CCGATTCTCGGTAACCTGCT
6588	*dnaN*	Reverse	CCATCCACCAGCTTCGAGGT
6585	*hemD*	Forward	CCACTGATTGAATTTGTCGC
2172	*hemD*	Reverse	GTTGTCGGCGTTATCAGCGAC
6076	*hisD*	Forward	GTCGGTCTGTATATTCCCGG
6077	*hisD*	Reverse	GGTAATCGCATCCACCAAATC
2179	*purE*	Forward	CGCATTATTCCGGCGCGTGT
2180	*purE*	Reverse	CGCGGATCGGGATTTTCCAG
2191	*sucA*	Forward	AGCACCGAAGAGAAACGCTG
6594	*sucA*	Reverse	GGTTGTTGATAACCGATACGTAC
6591	*thrA*	Forward	ATCCCGGCCGATCACATGAT
6592	*thrA*	Reverse	CTCCAGCAGCCCCTCTTTCAG

50x TAE buffer

2 M Tris-acetate and 50 mM EDTA (pH 8.0)

20x PBS buffer

2.8 M NaCl, 30 mM KH_2PO_4, 162 mM $Na_2HPO_4.2H_2O$ (pH 7.4)

Agarose gel loading buffer

0.25% bromophenol blue and 40% (w/v) sucrose in distilled water.

2.5 Glass beads for storage of bacterial isolates

2.8-3.3 mm glass beads were washed in tap water with a commercial detergent. They were then soaked in 0.05 M HCl for five minutes to neutralize alkalinity. The glass beads were washed several times in tap water until the pH of the wash water was that of tap water. The beads are washed twice in distilled water and dried in an oven at 50°C. Approximately 120 glass beads are dispensed per 10 ml loosely capped culture tubes which are sterilized by autoclaving at 121°C for 20 min. The cryo-vials were filled with approximately 20 sterilized glass beads in a laminar air-flow workbench.

2.6 Software and databases

The computer programs used in the study are summarized in Table 2.10 and the databases in Table 2.11.

Table 2.10. Computer programs

Name	Source	Web address
Access 2002	Microsoft	office.microsoft.com/en-us/access/default.aspx
Bionumerics 4.5	Applied Maths, Sint-Martens-Latem, Belgium	www.applied-maths.com
ClonalFrame 1.1	Didelot, X. and Falush,D., University of Oxford, UK	www2.warwick.ac.uk/fac/sci/statistics/staff/research/didelot/clonalframe
DnaSP 4.0	Rozas,J., Librado,P., Sánchez-DelBarrio,J.C., Messeguer, X., Rozas,R., Universitat de Barcelona, Spain	www.ub.es/dnasp
eBURST 3.0	Spratt, B.G., Hanage, W.P., Li, B., Aanensen, D.M., Feil, E.J., Imperial College, London, U.K.	http://eburst.mlst.net
GenAlEx 6.0	Peakall, R. and Smouse,P., The Australian National University, Canberra, Australia	www.anu.edu.au/BoZo/GenAlEx
Illustrator CS2	Adobe Systems Inc., San Jose, California, USA	www.adobe.com
MEGA 4.0	Tamura,K., Dudley,J., Nei,M., Kumar,S., Arizona State University, Tempe, Arizona, USA	www.megasoftware.net/fixed_bugs.html
Modeltest 3.7	David Posada, University of Vigo, Spain	http://darwin.uvigo.es/software/modeltest.html
Origin 6.0	Microcal Software Inc., Northhampton, Massachusetts, USA	www.microcal.com
PAML 4.0	Yang,Z., Department of Biology, University College London, London, UK	abacus.gene.ucl.ac.uk/software/paml.html
PAST 1.73	Hammer,Ø., Paleontological Museum, University of Oslo, Oslo, Norway	folk.uio.no/ohammer/past/download.html
Paup 4b10	Sinauer Associates Inc. Publishers, Sunderland, Massachusetts, USA	paup.csit.fsu.edu/

Table 2.10 continued…

Name	Source	Web address
RDP 3.0	Martin,D., University of Cape Town, Cape Town, South Africa	darwin.uvigo.es/rdp/rdp.html
Reference manager 10.0	ISI ResearchSoft, Carlsbad, California, USA	www.refman.com
Splitstree 4.0	Hudson,D.H. and Bryant,D., Tuebingen University, Tuebingen, Germany	www.splitstree.org
STATISTICA	StatSoft, In., Tulsa, U.S.A.	www.statsoft.com
UniFrac	Lozupone,C., Hamady,M., Knight,R., University of Colorado at Boulder, Boulder, USA	bmf2.colorado.edu/unifrac/index.psp

Table 2.11. Databases

Database	Web address
PubMed	www.ncbi.nlm.nih.gov/sites/entrez?db=pubmed
Salmonella MLST	web.mpiib-berlin.mpg.de/mlst/dbs/Senterica

3. Evolution and population structure of Newport

3.1 Introduction

Salmonellosis continues to be a major global public health concern due to a high number of human cases every year (World Health Organization, 2005). Newport has emerged as one of the most common serovars isolated from both humans and food animals in recent years (Centers for Disease Control and Prevention, 2002; Animal Disease Diagnostic Laboratory, 2004; Poppe *et al.*, 2006). It is the third most frequent serovar in the United States (Centers for Disease Control and Prevention, 2006) and one of the top 15 serovars which infects humans in Europe (European Centre for Disease Prevention and Control, 2008). Multidrug resistant Newport isolates, especially those resistant to expanded-spectrum cephalosporins (MDR-AmpC), have become a serious problem among food animals which are a major source of human infections (Zhao *et al.*, 2001; Gupta *et al.*, 2003; Devasia *et al.*, 2005; Poppe *et al.*, 2006; Egorova *et al.*, 2008). Outbreaks of MDR-AmpC Newport have been reported in the U.S.A. and France (Centers for Disease Control and Prevention, 2002; Espie *et al.*, 2005). MDR-AmpC Newport strains are resistant to ampicillin, chloramphenicol, streptomycin, sulphamethoxazole, tetracycline, amoxicillin/clavulanic acid, cephalothin, cefoxitin and ceftiofur and have a decreased susceptibility to ceftriaxone (Centers for Disease Control and Prevention, 2002; Devasia *et al.*, 2005).

The population structure of some serovars of subspecies *enterica* is monophyletic while others are polyphyletic (Beltran *et al.*, 1988). Newport is polyphyletic according to both MLEE and MLST studies (Beltran *et al.*, 1988; Sukhnanand *et al.*, 2005; Torpdahl *et al.*, 2005; Harbottle *et al.*, 2006). Two distantly related lineages were identified for Newport, one associated with humans and the second with domesticated animals (Beltran *et al.*, 1988; Sukhnanand *et al.*, 2005). However, the population structure of Newport has not yet been examined in detail. I analyzed 384 Newport isolates in order to study the evolution and population structure of this serovar using an MLST scheme (Torpdahl *et al.*, 2005). For comparative analyses, MLST data of four other serovars were also included, namely Enteritidis (including Gallinarum and Pullorum, avian-adapted variants), Kentucky, Paratyphi B (including *d*-tartrate positive variant Paratyphi B var. Java) and Typhimurium. These serovars were chosen because MLST data for more than fifty isolates are available from each of these serovars.

3.2 Methods

3.2.1 Bacterial isolates

149 Newport isolates were obtained from various *Salmonella* reference laboratories in Europe: 79 isolates were provided by Dr. Francois-Xavier Weill, Pasteur Institute, Paris, France; 50 by Dr. Wolfgang Rabsch, Robert Koch Institute, Wernigerode, Germany; and 20 by Dr. Reiner Helmuth, National Salmonella Reference Laboratory, Federal Institute for Risk Assessment (BFR), Berlin, Germany.

In addition, three isolates were obtained from Dr. Steffen Porwollik, Sidney Kimmel Cancer Center, San Diego, U.S.A and one isolate from Prof. Roy Curtiss III, Center for Infectious Diseases and Vaccinology, The Biodesign Institute, Arizona State University, Tempe, U.S.A. Finally, three Newport isolates from the SARB collection that had been tested by MLEE (SARB36, SARB37 and SARB38) (Boyd *et al.*, 1993) were obtained from Dr. Fidelma Boyd, University College Cork, Ireland.

MLST data for 135 Newport isolates was provided by Dr. Heather Harbottle, Center for Veterinary Medicine, U. S. Food and Drug Administration, Laurel, USA. Additional data were from the published literature (Harbottle *et al.*, 2006; Torpdahl *et al.*, 2005) and data for nine other isolates were from the *Salmonella* MLST website (http://web.mpiib-berlin.mpg.de/mlst/dbs/Senterica).

In all, data were available for 384 Newport isolates (Table 3.1). MLST data for serovars Enteritidis, Kentucky, Paratyphi B and Typhimurium were downloaded from the *Salmonella* MLST website (http://web.mpiib-berlin.mpg.de/mlst/dbs/Senterica).

3.2.2 Cultivation and storage of isolates

All bacterial isolates were streaked on LB agar plates and grown for 16-18 h at 37°C. A single colony was suspended in 100 µl of LB broth and spread on a fresh LB agar plate. After overnight growth, one third of the bacterial lawn was suspended in 1 ml of 10% skimmed milk. The suspension was added to a cryo-vial containing approximately 20 sterilized glass beads and was gently mixed by pipetting, resulting in a homogenous suspension. Excess fluid was discarded and the vials were stored at -80°C.

Table 3.1. Sources of Newport isolates

Year of Isolation (No. of isolates)	Host	Continent		
		Europe	North America	Others[a]
1940-1959 (33)[b]	Human	6	4	10
	Rat			1
	Reptile			7
	Swine	1		1
	Unknown		3	
1960-1979 (117)	Chicken			1
	Frog legs	3		
	Equine	1	1	
	Human	30	77	2
	Lion			1
	Swine			1
1980-2005 (226)	Bovine		20	
	Chicken	2	10	
	Food/feed	4	9	
	Human	60	75	
	Reptile	20	1	
	Swine		13	
	Turkey		9	
	Unknown	3		
Unknown (8)	Human	2	2	
	Swine		1	
	Unknown			3
Total		132	225	27

[a] Isolates from Africa, Asia, South America or of unknown sources.

[b] One strain was isolated in 1918.

3.2.3 Antimicrobial susceptibility typing

Disc diffusion and broth microdilution methods were used for antimicrobial susceptibility typing of the Newport isolates as follows:

3.2.3.1 Disc diffusion method

Newport isolates were tested for resistance to amikacin, cephalothin, cefoxitin and ceftriaxone by the disc diffusion method. Bacterial cultures were streaked to LB agar plates and grown overnight at 37ºC. Four ml of LB-broth was inoculated with a loop full of bacteria from these plates and incubated at 37ºC for 18 h. 100 µl of a 1:100 dilution was spread on a Müller-Hinton agar plate (Oxoid, UK) and 6 mm discs containing antimicrobials (Oxoid, UK) were placed on the agar surface. These plates were incubated for 18 h at 37ºC. Zone diameters were measured and compared with break points that were recommended by the CLSI (NCCLS, 2002).

3.2.3.2 Broth microdilution method

A total of twelve antimicrobials were tested by the broth microdilution method, namely ampicillin, amoxicillin-clavulanic acid, ceftiofur, chloramphenicol, ciprofloxacin, gentamicin, kanamycin, nalidixic acid, streptomycin, sulfamethoxazole, tetracycline and trimethoprim-sulfamethoxazole. Bacterial cultures were streaked on Müller-Hinton agar plates and incubated at 37ºC for 18 h. Bacteria from these plates were suspended in sterile normal saline solution (0.85% NaCl) to a turbidity of 0.5 McFarland using the Sensititer System (Autoinoculator INO2, Trek Diagnostic Systems Ltd., UK). 15 µl of this suspension was mixed with 11 ml Müller-Hinton-broth (Oxoid, UK). 100 µl of the inoculated broth was dispensed to each well of a microtiter plate (NLMV1A, Trek Diagnostic Systmes Ltd., UK) that contained lyophilized antimicrobial agents in different concentrations. The plates were sealed with sealing films and incubated at 37ºC for 18 h. The plates were read with a semi-automatic Sensitouch System (Accumed International Ltd., UK) in order to determine the minimum inhibitory concentration (MIC µg/ml) as per the CLSI guidelines (NCCLS, 2002).

3.2.4 DNA isolation

DNA was extracted from liquid cultures grown overnight at 37°C in LB broth using the JETFLEX Genomic DNA purification kit (GENOMED). Approximately 1.5 ml of LB broth was added to each well in a 96 deep-well plate and inoculated with one glass-bead from the -80°C bacterial stock. The cultures were grown overnight in a shaking incubator at 37°C and 80-100 rpm. The plate was centrifuged at 3,220×g' for 5 min and the supernatant was discarded. The pellets were dissolved in 300 µl of cell lysis buffer and 10 µl of proteinase K (20 mg/ml) was added to each well followed by incubation at 58°C for 2 h in a waterbath. After incubation, 10 µl of RNase (4 mg/ml) was added to each well and the plate was incubated at 37°C for 10 min in an incubator. 150 µl of protein precipitation buffer and 50 µl of pellet compactor were added to each well and mixed thoroughly to precipitate the cell debris. The plate was centrifuged at 3,220×g' for 15 min, vortexed and incubated on ice for 10 min before re-centrifugation at 3,220×g' for 15 min.

The clear supernatant was transferred to a second 96 deep-well plate and 450 µl of iso-propanol was added to each well. The contents were mixed by gently inverting the plate after sealing with a sealing film. The plate was incubated on ice for 15 min followed by centrifugation at 3,220×g' for 10 min. The supernatant was discarded by inverting the plate

on a paper towel and 500 µl of 70% ethanol (room temperature) was added to wash the pellet. The plate was centrifuged at 3,220×g' for 5 min and the supernatant was discarded in the same manner. The washing with 70% ethanol (room temperature) was repeated once again. The pellets were air dried and 200 µl of Tris buffer pH 8.0 (10 mM) was added to each of the wells. The plate was sealed with sealing film and left at 4°C overnight to dissolve the DNA.

Five µl of each DNA was diluted 1:10 with Tris buffer pH 8.0 (10mM) in a microtitre plate and quantified using the spectrophotometer SPECTRA-MAX 190 or GeneQuant at the wavelengths 260 and 280 nm. One unit of absorbance of dsDNA at 260 nm is equal to 50 ng/µl. The OD at 280 is used to check the quality of DNA. The ratio of ODs should be ~1.8. A ratio of < 1.7 indicates contamination with protein or other reagents left due to improper washing and ≥ 2.0 indicates RNA contamination. DNA whose A_{260}/A_{280} OD ratio was < 1.7 ratio was re-purified by repeating the entire procedure after adding 300 µl cell lysis buffer and 10 µl proteinase K (20 mg/ml). 10 µl of RNase (4 mg/ml) was added to DNA with A_{260}/A_{280} ratio of > 2.0 followed by incubation at 37°C for 10 min in a waterbath. The DNA was precipitated by adding 1 volume iso-propanol and washed with 70% ethanol (room temperature) as described. DNAs with A_{260}/A_{280} ratio between 1.7-1.9 were diluted to the concentration of 50 ng/µl with Tris buffer pH 8.0 (10 mM) and were stored at -20°C. 20 µl of DNA from each of these stocks was further diluted to 5 ng/µl by adding 180 µl Tris buffer pH 8.0 (10 mM) which was stored at 4°C and was used for PCR amplification.

3.2.5 PCR amplification and sequencing

3.2.5.1 PCR amplification

Fragments of seven housekeeping genes were amplified as previously described (Harbottle *et al.*, 2006; Torpdahl *et al.*, 2005). The primers used for amplification are listed in Table 2.9 of Chapter 2. The reaction mix contained 1x reaction buffer with $MgCl_2$, 200 µM deoxynucleoside triphosphates (dNTPs), 10 pmol of each primer, 1U of *Taq* polymerase and 10 ng of template DNA in 15 µl of reaction volume. PCR cycling conditions were as follows:

Step 1. 94°C : 5 min

(35 cycles from step 2 to 4)

Step 2. 94°C : 1 min

Step 3. 55°C : 1 min

Step 4. 72°C : 1 min

Step 5. 72°C : 5 min

Step 6. 4°C : ∞

3.2.5.2 Agarose gel electrophoresis

One gram of agarose was dissolved in 100 ml of TAE buffer in a microwave oven. The solution was allowed to cool to 50-60°C and 5 µl of 10 mg/ml ethidium bromide solution was added. The gel was poured in a 13 X 13.5 cm plate and four combs were inserted, each with 26 teeth. Once the gel had solidified, the combs were removed and 3 µl of each PCR product plus 1 µl of bromophenol blue dye were loaded. The DNA was resolved by electrophoresis at 5V/cm until the loading dye had reached the beginning of the second block. The gel documentation system BIO-RAD Gel Doc 2000 was used for visualization of amplified products and photography.

3.2.5.3 Cleaning of amplified products

Amplified products were treated with shrimp alkaline phosphatase (SAP) and exonuclease I (ExoI) to remove unused dNTPs and primers. 0.5 µl of ExoI (10 U/µl) and 0.5 µl of SAP (1 U/µl) were added to 12 µl of PCR product in a 96 well plate. The plate was sealed with adhesive sealing film and incubated as follows:

Step 1. 37°C : 60 min

Step 2. 80°C : 15 min

Step 3. 4°C : ∞

3.2.5.4 Sequencing of purified PCR products

PCR products were diluted with sterile distilled water to a concentration of 5 ng/µl. The primers used for sequencing are listed in Table 2.9 of Chapter 2. Sequencing reaction mix contained 1.875 µl 5X sequencing buffer, 0.25 µl Big Dye Terminator v3.0 and 3 pmol of primer. 2.0 µl of diluted PCR product was used as template and sterile distilled water was added to make up the reaction volume to 10 µl. The conditions used are as follows:

Step 1. 96°C : 2 min

(30 cycles from step 2 to 4)

Step 2. 96°C : 10 sec
Step 3. 50°C : 5 sec
Step 4. 60°C : 2 min
Step 5. 4°C : ∞

The sequenced products were precipitated by ethanol precipitation. 10 µl of sterile distilled water was added to 10 µl of sequence reaction. 7 ml absolute ethyl-alcohol was mixed with 280 µl of 3 M sodium-acetate, pH 4.6 and 52 µl of the mix was added to each well. The microtitre plate was sealed with sealing film, briefly vortexed to mix the content and incubated at room temperature for 45 min.

The plate was centrifuged at 2,750×g' for 1 h at 4^{o}C and the supernatant was discarded by gently inverting onto a paper towel followed by an inverted spin onto a piece of Whatman 3MM paper at 500×g' for 1 min. DNA pellets were washed twice by adding 150 µl of 70% ethyl-alcohol (room temperature) and centifugation at 2,750×g' for 10 min. The supernatant was discarded as before. Dried pellets were dissolved in 10 µl of Hi Di formamide and sequenced with an ABI PRISM DNA analyzer 3770 (Applied Biosystems, Foster City, Calif).

3.2.6 MLST analyses

Strain information that includes strain designation, subspecies, serovar, year of isolation, host, continent and source were stored in a Bionumerics 4.5 database (Applied Maths, Sint-Martens-Latem, Belgium). The nucleotide sequences of gene fragments were also stored in the same database. They were assembled and trimmed using scripts written by Prof. Mark Achtman. New alleles and STs were submitted to the *Salmonella* MLST website for the assignment of allele and ST numbers after curation. Final allele and ST designations were assigned to the entries in the Bionumerics 4.5 database using scripts also written by Prof. Mark Achtman. Nucleotide sequences of gene fragments were concatenated in the order *aroC*, *dnaN*, *hemD*, *hisD*, *purE*, *sucA* and *thrA* for analyses of concatenated sequence alignments. Alignments of gene fragments and concatenated sequences were exported from Bionumerics 4.5 for further analyses.

3.2.7 Evolutionary analyses

3.2.7.1 Clonal Frame

The computer program ClonalFrame (Didelot and Falush, 2006) was used to infer clonal relationships among strains on the basis of multilocus sequence data. ClonalFrame is based on a coalescent approach which assumes that the bacteria in a sample come from a constant-sized population and all individuals are equally likely to reproduce. The program identifies potential regions within nucleotide sequences that have been imported from other bacteria and calculate genealogies after ignoring these imports which minimizes the effects of recombination. A consensus tree based on the posterior probabilities of genealogies is generated, depicting clonal relationships among the strains.

Ten ClonalFrame runs were computed on unique STs from the entire dataset, each with 100,000 iterations after 100,000 burn-in iterations. A 50% consensus tree was generated on the outputs of these 10 runs using the graphical user interface of the program.

3.2.7.2 Neighbor-net

Neighbor-net is a distance based method that constructs phylogenetic networks using nucleotide sequence data (Bryant and Moulton, 2004). An agglomerative process of pairing nodes is followed by collection of weighted splits which are converted into a phylogenetic network (Bryant and Moulton, 2004). The method is highly efficient for resolving conflicting signals in the dataset in comparison to other methods like split decomposition (Bryant and Moulton, 2004).

Various DNA substitution models were tested on an alignment of unique concatenated sequences in PAUP 4.0 (Swofford, 1998) and the model that fits best was chosen using Modeltest 3.7 (Posada and Crandall, 1998). This model and parameter estimates were used to generate a phylogenetic network on the dataset using Neighbor-net algorithm implemented in SplitsTree 4.0 (Huson and Bryant, 2006).

3.2.7.3 Minimal spanning tree (MS_{TREE})

The MS_{TREE} is a graphical tool that links the nodes by unique minimal paths in a given dataset i.e. total summed distance of all branches is minimized. The MST algorithm implemented in Bionumerics 4.5 was used to generate an MS_{TREE} from the allelic profiles of all isolates. The algorithm uses an ST with highest numbers of single locus variants (SLVs) as a root node and derives other STs from it. STs that shared up to three alleles were connected by crosslinks.

A second MS_{TREE} was generated from allelic profiles only for Newport isolates without crosslinks.

3.2.7.4 eBURST

eBURST is an algorithm that identifies groups of closely related sequence types from MLST data (Feil *et al.*, 2004). The algorithm attempts to predict a founder for each group based on the abundance of STs that are linked to a particular ST at a single allelic difference. If two STs in a group have the same number of SLVs, then the one with the larger number of double locus variants (DLVs) is chosen. The output is displayed in the form of a radial diagram where descendent genotypes (SLVs) are linked to the predicted founder at the centre. Other STs within a group (except for the predicted founder) that have ≥ 3 SLVs are predicted to be co-founders. It has been a popular method to study the evolutionary descent among isolates using MLST data (Enright *et al.*, 2002; Feil *et al.*, 2004; Honsa *et al.*, 2008; Feil *et al.*, 2003). The algorithm was tested on allelic data of Newport isolates using eBURST 3.0 (http://eburst.mlst.net).

3.2.7.5 Reticulate

The concatenated sequence alignments from each group identified by the ClonalFrame were independently tested for reticulate evolution by the program Reticulate (Jakobsen and Easteal, 1996). This program calculates the compatibility between pairs of parsimony informative sites in a given alignment, resulting in an overall compatibility score ranging from 0 (no compatibility) to 1 (highly compatible). A pair of informative site is compatible if $c = n–1$ for both sites, where c is the minimum number of changes that are required to construct a genealogy and n is the total number of polymorphic nucleotides. If $c > n–1$ for any one or both sites, they are incompatible. Such incompatibilities generally result from recurrent mutations, recombination or lateral gene transfer. An overall low compatibility score indicate that different nucleotides followed different evolutionary paths and a phylogenetic tree on the concatenated sequences would not reflect true evolutionary relatedness.

3.2.7.6 Pairwise homoplasy index (Φw) test of recombination

The Φw test is a program that detects recombination on the basis of low compatibility between informative sites in a sequence alignment (Bruen *et al.*, 2006). The program

calculates refined incompatibilities between the pairs of parsimony informative sites as per the following equation:

$i(\chi_i, \chi_j) = l(\chi_i, \chi_j) - (|\chi_i|-1) - (|\chi_j| -1)$

$i(\chi_i, \chi_j)$ = refined incompatibility score for parsimony informative sites i and j

$l(\chi_i, \chi_j)$ = minimum number of mutations required to represent genealogy by any tree for these sites

$|\chi_i|$ = number of different states at site i

$|\chi_j|$ = number of different states at site j

The mean of refined incompatibility is calculated for nearby sites in the diagonal for the first k rows of the incompatibility matrix as the pairwise homoplasy index or Φw statistics using the following equation [taken from Bruen et al., 2006].

$$\Phi w = \frac{2}{k(2n-k-1)} \sum_{j=1}^{k} \sum_{i=1}^{n-j} i(\chi_i, \chi_{i+j})$$

k = wq

w = width of the window in basepairs (set to 100 by default)

q = proportion of parsimony informative sites within the alignment

n = number of total informative sites

k(2n-k-1)/2 = normalizing factor

The significance of the observed Φw statistics is obtained by permutation test. The concatenated sequence alignments of ClonalFrame groups were tested for recombination by Φw test.

3.2.8 Characteristics of housekeeping genes

The average pairwise nucleotide diversity per site (π) with Jukes-Cantor correction was calculated for gene alignments in each group using MEGA version 4.0 (Tamura *et al.*, 2007).

$$\pi' = \frac{1}{[n(n-1)/2]} \sum_{i<j} \pi_{ij}$$ [taken from (Page and Holmes, 1998)]

π' is the average pairwise nucleotide diversity, n is the number of sequences in the sample and π_{ij} is the difference between the i^{th} and the j^{th} sequence. π can be calculated by dividing π' by the total length of the sequences.

Assuming that substitutions at all the sites are equally likely and that the frequencies of all four bases are the same, Jukes and Cantor proposed that the mean number of base differences (d) between a pair of sequences is related to the proportion of different nucleotides (p) by

$d = - ¾ \ln (1 – 4/3\ p)$ [taken from (Page and Holmes, 1998)]

The mean of non-synonymous substitutions per non-synonymous site (Ka) and synonymous substitutions per synonymous site (Ks) were computed for gene alignments using DnaSP version 4.0 (Rozas *et al.*, 2003). Substitutions in a protein coding sequence that result in amino acid changes are known as non-synonymous substitutions while those that does not change the amino acid are called synonymous. The Ka/Ks ratio is commonly used to estimate selection pressure on a protein coding sequence. A value of Ka/Ks > 1.0 indicates diversifying selection whereas Ka/Ks = 1.0 indicates neutral evolution and Ka/Ks < 1.0 indicates purifying selection.

3.3 Results

3.3.1 Population structure of Newport and other serovars

An alignment of unique concatenated sequences was tested by ClonalFrame to infer their phylogenetic relatedness. I assigned each cluster of ≥ 3 STs or a distinct ST with ≥ 10 isolates as a group and STs that did not meet these criteria were treated as singletons. The STs of serovar Newport grouped into three distinct clusters that were designated Newport-I, Newport-II and Newport-III (Fig. 3.1). Newport-I was a group of three STs that contained 8% of the Newport isolates (Table 3.2). Newport-II and Newport-III were more diverse and more common, each containing 47% of the Newport STs and 58% and 34% of the isolates, respectively. STs, allelic profiles and the number of isolates within each Newport group are summarized in Table 3.3.

Paratyphi B (including var. Java) is a highly heterogeneous serovar because four distinct groups were identified among the 18 STs that were identified among 66 isolates. Kentucky is also a polyphyletic serovar because five STs were assigned to two groups, designated as

Kentucky-I and Kentucky-II. In contrast, most of the Enteritidis and Typhimurium STs clustered into a single group for each serovar. However, two Enteritidis STs (ST6 and ST77) and two Typhimurium STs (ST36 and ST207) were singletons. Each of the singleton STs was distant from the main groups and contained only one to two isolates.

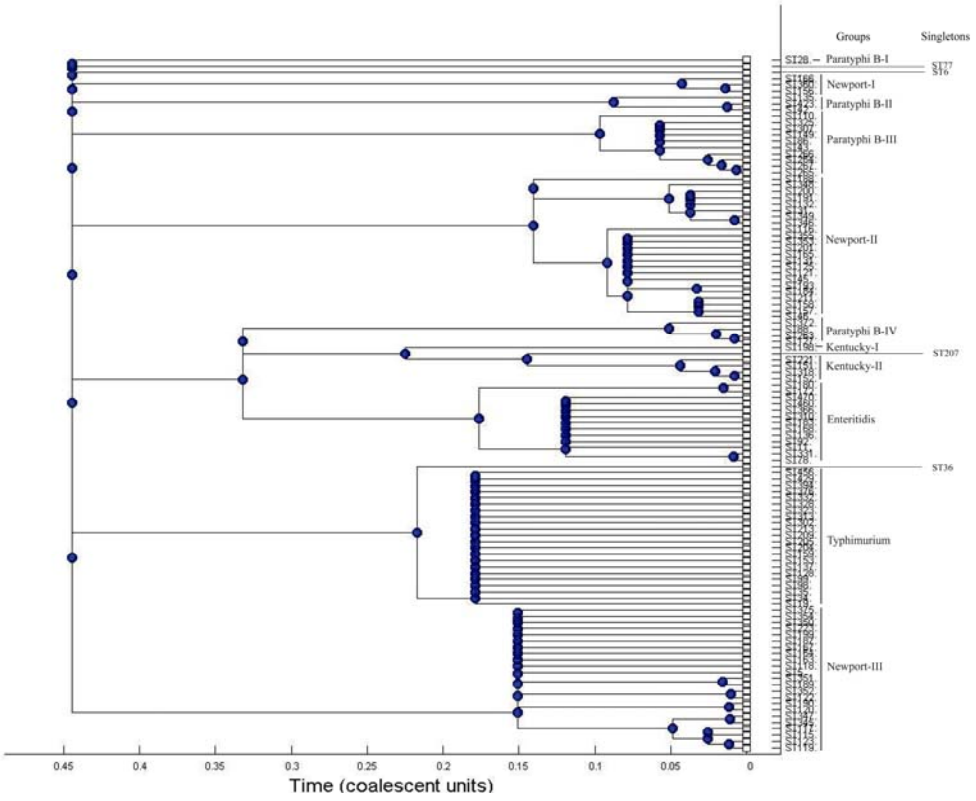

Fig. 3.1. A 50% consensus of 10 evolutionary trees generated using ClonalFrame 1.1.

Serovars of subspecies *enterica* vary greatly in their population structures and can contain between one to four distinct groups per serovar. The degree of polyphyly within serovar Newport was higher than Enteritidis, Typhimurium or Kentucky but lower than Paratyphi B.

Recombination has been an important source of variation among housekeeping genes within subspecies *enterica* (Brown *et al.*, 2003; Octavia and Lan, 2006; Falush *et al.*, 2006). The effects of recombination are minimized when ClonalFrame is used to deduce genealogies. Therefore, to test whether the structures of the groups identified by ClonalFrame are different when recombination is not excluded, I used the Neighbor-net algorithm on the alignment of unique concatenated sequences from all serovars. The conflicting signals introduced to nucleotide sequences by recombination or recurrent mutations are displayed as parallel paths in a phylogenetic network by Neighbor-net.

Table 3.2. Groups identified using ClonalFrame for each serovar

Serotype	No. of isolates	ClonalFrame groups	No. of STs	STs
Enteritidis[1]	141	Enteritidis	13	11, 470, 183, 366, 78, 310, 168, 136, 460, 92, 331, 180, 172
	2	Singletons	2	6, 77
Kentucky	13	Kentucky-I	1	198
	159	Kentucky-II	4	151,152, 221, 318
Newport	32	Newport-I	3	156, 166, 360
	222	Newport-II	23	31, 132, 348, 188, 191, 200, 346, 349, 193, 45, 116, 121, 125, 131, 165, 353, 355, 46, 157, 211, 158, 201, 184
	130	Newport-III	23	118, 189, 122, 199, 164, 163, 345, 351, 120, 190, 223, 167, 5, 187, 347, 352, 354, 115, 119, 117, 375, 123, 350
Paratyphi B[2]	16	Paratyphi B-I	1	28
	5	Paratyphi B-II	3	42, 423, 135
	29	Paratyphi B-III	10	86, 43, 267, 266, 265, 264, 149, 307, 110, 325
	16	Paratyphi B-IV	4	88, 127, 263, 372
Typhimurium	362	Typhimurium	22	19, 128, 376, 209, 205, 204, 159, 137, 429, 313, 35, 99, 456, 153, 213, 302, 98, 323, 332, 328, 34, 394
	3	Singleton	2	36, 207
Total	1130		111	

[1]Isolates of Gallinarum and Pullorum are also included

[2]Isolates of Paratyphi B var. Java are also included

Neighbor-net identified similar groups to those defined by ClonalFrame, both in numbers and composition (Fig. 3.2). However, Neighbor-net separated three STs from Newport-II and one from Paratyphi B-II, which might indicate that multiple alleles in these STs were acquired by homologous recombination. Most STs in Enteritidis and Typhimurium were arranged in a radial symmetry. However, conflicting signals were resolved for multiple STs in the other groups as indicated by parallel paths within these groups.

In agreement with the conclusions based on the ClonalFrame output, a serovar to serovar variation in population structures was also observed within subspecies *enterica* with Neighbor-net. The radial symmetry in Enteritidis and Typhimurium might indicate that most diversity in these groups was generated by mutations whereas parallel paths for multiple STs in other groups might be indicative of substantial recombination.

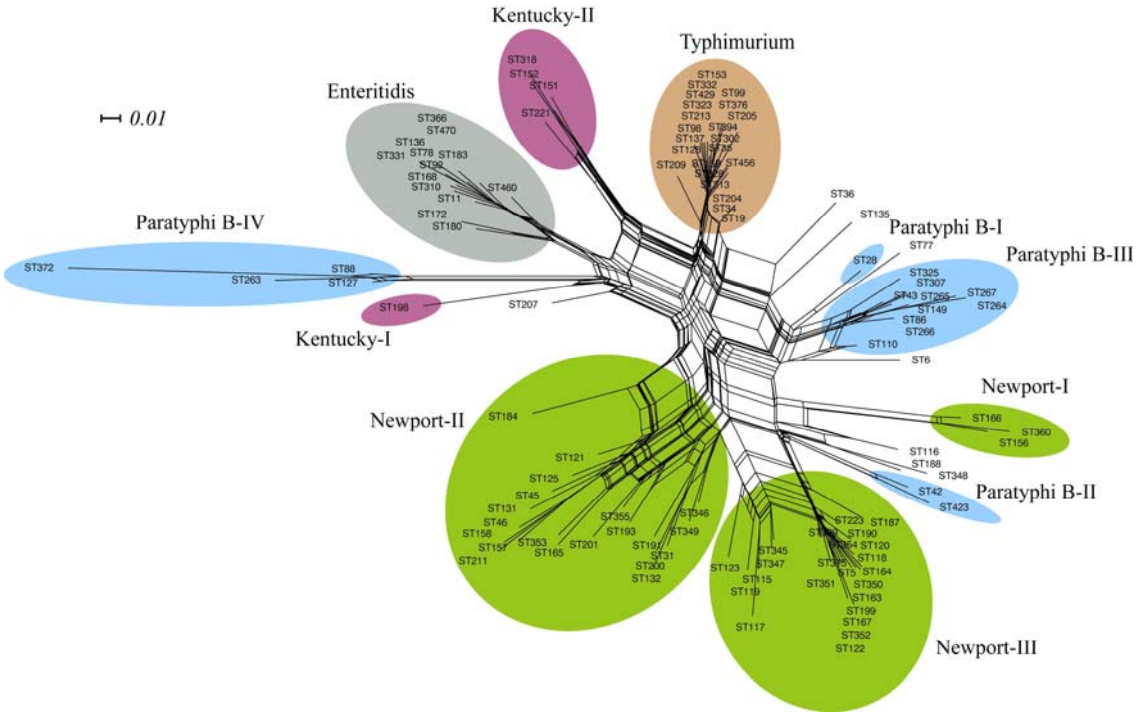

Fig. 3.2. Neighbour-net on concatenated sequence alignment following GTR+I substitution model and parameter estimates chosen using Modeltest 3.7

Allele based phylogeny has been used extensively to study bacterial population structure and identify the groups of closely related genotypes. To compare the sequence based phylogeny with the groups based on shared alleles, an MS_{TREE} was generated from the allelic profiles of isolates. Groups of at least 3 STs that shared six of the seven alleles with one another were assigned a group designation. Distinct STs with ≥ 10 isolates were also designated as groups whereas those with < 10 isolates were treated as singletons. STs that shared up to three alleles were recognized by the use of cross-linking in order to help to resolve relatedness between the STs of different serovars up to the level of four allelic differences (Fig. 3.3).

The numbers and compositions of the groups were again consistent with the ClonalFrame results for each serovar. However, Newport-II was divided into two subgroups and four singletons STs. All singletons were linked to ST45 in one of the subgroups at a distance of two alleles (DLVs). One of these singletons, ST193, connected both the subgroups because it also shared five alleles with ST31 of the second subgroup (Fig. 3.3, box a). However, the structure of Newport-II was identical with the ClonalFrame output when the algorithm was allowed to identify groups based on sharing of five alleles instead of six (Fig. 3.3, box b).

Moreover, numerous cross-links between the STs from both the subgroups suggest that they are closely related. Similarly, two STs (ST172 and ST180) from Enteritidis, one (ST135) from Paratyphi B-II and two (ST123 and ST350) from Newport-III were also separated. These STs were also linked to the respective groups by multiple cross-links and lacked close associations with other groups. Therefore, the population structures of all the serovars were consistent with the ClonalFrame and Neighbor-net results, suggesting that the serovar Newport was more heterogeneous than Enteritidis, Kentucky or Typhimurium but lesser than Paratyphi B.

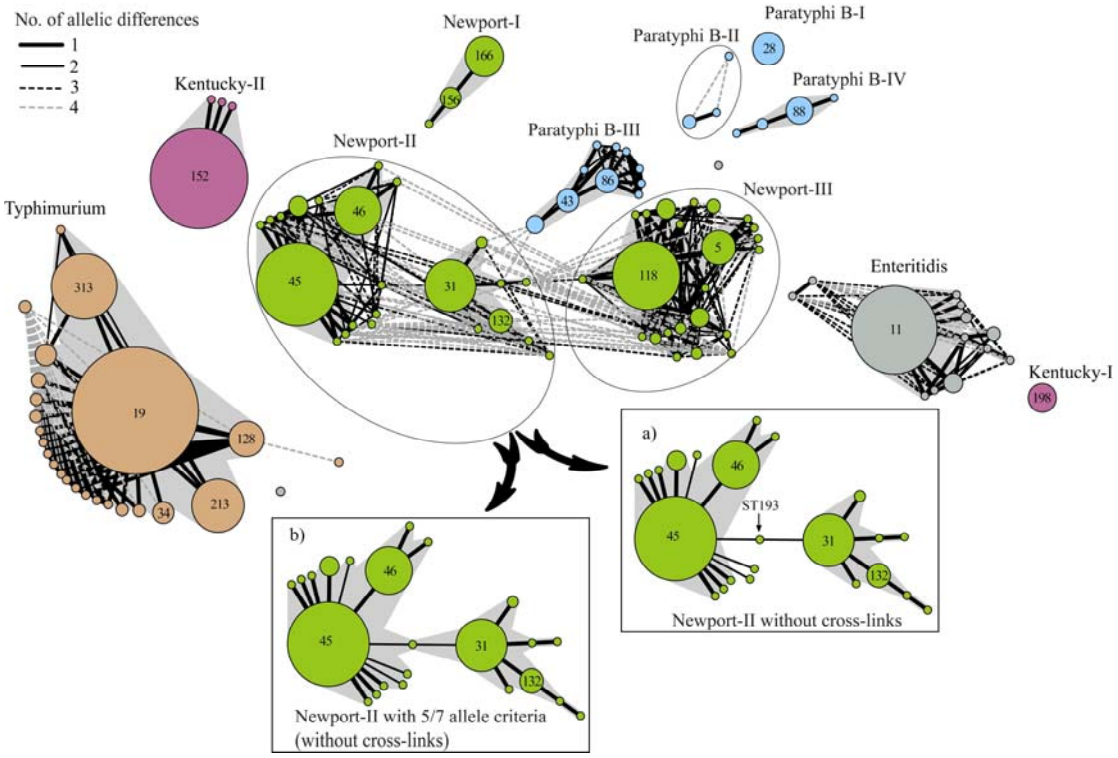

Figure 3.3 An MS_{TREE} on the allelic profiles of isolates with a crosslink distance of four alleles. Box a). the structure of Newport-II without cross-links and Box b). Newport-II when the groups of ≥ 3 STs that shared at least five of the seven alleles with at least one other ST in the group were considered.

Cross-links were also observed between several STs of Newport-II and Newport-III, indicating that multiple STs between these groups share up to 4 alleles. One alternative is that these two groups are merging together due to frequent homologous recombination between the groups. A second is that they are drifting apart by recombining with other serovars.

Table 3.3. Allelic profiles in 384 Newport isolates

Groupings	ST	No. of Strains	*aroC*	*dnaN*	*hemD*	*hisD*	*purE*	*sucA*	*thrA*
Newport-I	166	24	5	14	6	12	5	14	58
	156	7	63	14	6	12	5	14	58
	360	1	63	14	6	12	5	14	12
Newport-II	45	109	10	7	21	14	15	12	12
	31	43	2	2	15	14	15	20	12
	46	36	10	7	21	12	15	12	12
	132	9	2	57	15	14	15	20	12
	116	6	10	7	21	14	15	39	12
	191	2	2	2	15	14	15	20	42
	121	1	10	7	21	14	36	12	12
	125	1	10	7	21	14	15	53	12
	131	1	10	56	21	14	15	12	12
	157	1	10	61	21	12	15	12	12
	158	1	10	62	21	67	15	12	12
	165	1	10	7	21	14	58	12	12
	184	1	10	64	21	14	15	12	33
	188	1	2	7	15	14	15	39	12
	193	1	10	2	21	14	15	20	12
	200	1	77	2	15	14	15	20	12
	201	1	16	7	21	14	15	12	42
	211	1	10	7	21	12	15	12	71
	346	1	2	2	15	14	36	20	12
	348	1	2	2	15	14	15	39	12
	349	1	2	2	15	43	36	20	12
	353	1	10	7	21	43	15	12	12
	355	1	10	7	21	14	15	7	12
Newport-III	118	75	16	2	45	43	36	39	42
	5	19	16	43	45	43	36	39	42
	115	6	16	2	45	43	36	12	42
	164	6	16	2	45	70	36	39	42
	350	3	16	2	95	117	36	39	117
	167	2	16	2	61	71	36	39	42
	189	2	16	2	45	14	36	39	42
	223	2	16	2	61	43	36	39	42
	117	1	16	2	45	43	15	12	42
	119	1	16	2	18	43	36	12	42
	120	1	16	2	40	43	36	39	42
	122	1	16	2	45	43	15	39	42
	123	1	10	2	21	43	36	12	42
	163	1	16	2	60	43	36	39	42
	187	1	16	43	45	43	36	39	12
	190	1	16	2	40	43	36	19	42
	199	1	16	2	45	43	65	39	42
	345	1	16	2	45	43	36	20	42
	347	1	16	43	45	43	36	20	42
	351	1	16	2	45	62	36	39	42
	352	1	16	43	45	43	15	39	42
	354	1	16	43	45	43	71	39	42
	375	1	16	2	45	43	5	39	42

Newport is particularly different to other serovars. The groups of other serovars were more distinct because ≥ 5 alleles were different between the STs of any two groups within those serovars. However, an ST of Paratyphi B-III (ST110) shared three alleles with three STs

(ST31, ST191 and ST346) of Newport-II. It might indicate the genetic exchange between the isolates of different serovars within subspecies *enterica* or the remnants of the common ancestry. These results are indicative of inter-serovar recombination within subspecies *enterica*.

3.3.2 Properties of housekeeping genes in various groups

The sequence diversity at selectively neutral loci is the most appropriate to study the evolutionary relatedness among strains (Maiden *et al.*, 1998) whereas genes under positive selection might not correlate with the evolutionary history. To test the selective pressure on gene fragments, the Ka/Ks ratios were calculated for each fragment from all serovars. Most gene fragments were under strong purifying selection as indicated by a Ka/Ks ratio of << 1.0 except for *hemD* (Table 3.4) (Perez-Losada *et al.*, 2006). Although the value of Ka/Ks was <1.0 for the fragments of *hemD* indicating an absence of diversifying selection, it was comparatively much higher than other gene fragments. It indicates that the frequency of fixation of non-synonymous substitutions is higher in *hemD* than other gene fragments.

Table 3.4. Average pairwise distance at non-synonymous (Ka) and synonymous (Ks) sites among gene fragments

Gene fragment	Ka	Ks	Ka/Ks
aroC	0.00002	0.02364	0.0008
dnaN	0.00136	0.02271	0.0599
hemD	0.00522	0.02031	0.2570
hisD	0.00315	0.05104	0.0617
purE	0.00077	0.02613	0.0295
sucA	0.00009	0.02933	0.0031
thrA	0.00003	0.06131	0.0005

To compare the genetic diversity between the groups, the mean pairwise nucleotide distances were calculated for gene fragments within each group. Kentucky-I and Paratyphi B-I were single ST groups that did not have any polymorphic sites. Only ≤4 loci were polymorphic in Kentucky-II, Newport-I, Paratyphi B-II, Paratyphi B-III and Paratyphi B-IV (Table 3.5a) whereas ≥ 6 loci were polymorphic among remaining groups including Newport-II and III (Table 3.5b).

The numbers of polymorphic and informative sites were highly variable between Newport groups. Only two loci, *aroC* and *thrA*, were polymorphic in Newport-I, each has two alleles. In contrast, all the seven loci were polymorphic with 2-7 alleles in Newport-II and Newport-III. Nucleotide diversity on concatenated sequence alignment was significantly

lower within Newport-I than other Newport groups. The highest π value was observed for Newport-II whereas Newport-III showed intermediate genetic diversity between Newport-I and II.

Table 3.5a. Characteristics of gene fragments in groups with ≤ 4 polymorphic loci

ClonalFrame Group	Gene fragment	Alleles	Polymorphic sites	Informative sites	π
Kentucky-II	*hisD*	2	9	0	0.0002 ± 0.0001
(N = 159)	*sucA*	2	5	0	0.0001 ± 0.0001
	thrA	2	1	0	<0.0001 ± 0.0001
	Concatenated	4	15	0	0.0001 ± 0.0000
Newport-I	*aroC*	2	1	1	0.0008 ± 0.0008
(N = 32)	*thrA*	2	1	0	0.0001 ± 0.0001
	Concatenated	3	2	1	0.0001 ± 0.0001
ParatyphiB-II	*aroC*	2	5	0	0.0040 ± 0.0016
(N = 5)	*dnaN*	2	5	0	0.0040 ± 0.0018
	sucA	2	1	0	0.0008 ± 0.0007
	thrA	3	10	0	0.0081 ± 0.0021
	Concatenated	3	21	0	0.0025 ± 0.0005
ParatyphiB-III	*aroC*	5	16	7	0.0038 ± 0.0010
(N = 29)	*dnaN*	2	5	5	0.0030 ± 0.0014
	purE	4	6	1	0.0022 ± 0.0012
	thrA	2	1	1	0.0003 ± 0.0002
	Concatenated	10	28	14	0.0013 ± 0.0004
ParatyphiB-IV	*aroC*	2	8	0	0.0020 ± 0.0007
(N = 16)	*dnaN*	2	1	1	0.0006 ± 0.0006
	purE	2	26	0	0.0085 ± 0.0016
	Concatenated	4	35	1	0.0014 ± 0.0002

Note: 1. Monomorphic genes within each group were not mentioned.
2. Kentucky-I and Paratyphi B-I, single ST groups, were also not mentioned because all gene fragments were monomorphic.

Although only ≤ 4 loci were polymorphic in Newport-I and Kentucky-II while six or more loci were polymorphic in Enteritidis and Typhimurium, the π values were comparable between these groups. Similarly, nucleotide diversity was comparable between Newport-II and Paratyphi B-II even though all seven loci were polymorphic in Newport-II whereas three of the loci were monomorphic in Paratyphi B-II. The π value in Newport-III was not comparable to any other group because they were higher than Enteritidis, Kentucky-II and Typhimurium but lower than Paratyphi B-II, III and IV. Therefore, like the variation in the population structure, nucleotide diversity among gene fragments also varied greatly between the serovars as well as groups within a serovar (Table 3.5a & b). However, I did not observe any correlation between the two.

Table 3.5b. Characteristics of gene fragments in other groups

ClonalFrame Group	Gene fragment	Alleles	Polymorphic sites	Informative sites	π
Enteritidis	*aroC*	5	8	5	0.0004 ± 0.0001
(N = 141)	*dnaN*	1	0	0	<0.0003 ± 0.0003
	hemD	3	3	2	0.0002 ± 0.0001
	hisD	2	2	2	0.0001 ± 0.0001
	purE	5	5	4	0.0006 ± 0.0003
	sucA	2	1	1	0.0001 ± 0.0001
	thrA	3	7	4	0.0003 ± 0.0001
	Concatenated	13	26	18	0.0002 ± 0.0001
Newport-II	*aroC*	4	5	2	0.0016 ± 0.0010
(N = 222)	*dnaN*	7	11	6	0.0042 ± 0.0016
	hemD	2	6	6	0.0055 ± 0.0020
	hisD	4	11	10	0.0048 ± 0.0014
	purE	3	6	5	0.0004 ± 0.0001
	sucA	5	10	10	0.0042 ± 0.0017
	thrA	4	15	8	0.0006 ± 0.0002
	Concatenated	23	64	47	0.0031 ± 0.0005
Newport-III	*aroC*	2	3	0	0.0001 ± 0.0000
(N = 130)	*dnaN*	2	1	1	0.0006 ± 0.0006
	hemD	7	9	8	0.0009 ± 0.0003
	hisD	6	7	7	0.0007 ± 0.0003
	purE	5	7	5	0.0007 ± 0.0003
	sucA	4	10	10	0.0026 ± 0.0009
	thrA	3	9	1	0.0003 ± 0.0001
	Concatenated	23	46	32	0.0008 ± 0.0002
Typhimurium	*aroC*	3	2	2	<0.0001 ± 0.0001
(N = 364)	*dnaN*	2	1	1	0.0001 ± 0.0001
	hemD	3	2	2	0.0001 ± 0.0000
	hisD	2	1	0	<0.0001 ± 0.0001
	purE	4	3	3	0.0013 ± 0.0008
	sucA	9	13	3	0.0003 ± 0.0002
	thrA	5	7	0	0.0001 ± 0.0000
	Concatenated	23	29	10	0.0002 ± 0.0001

3.3.3 Source of diversity within each group

Parallel paths were observed in Neighbor-net possibly indicating higher frequencies of recombination than mutation except for Enteritidis and Typhimurium (Fig. 3.2). Therefore, the concatenated sequence alignment of each group was tested for recombination by Reticulate (Jakobsen and Easteal, 1996) and Φw test (Bruen *et al.*, 2006). These tests are based on the compatibilities between informative sites in a nucleotide sequence alignment which is considered to be one of the most powerful methods for recombination detection (Bruen *et al.*, 2006). Kentucky-I, Kentucky-II, Newport-I, Paratyphi B-I, Paratyphi B-II and Paratyphi B-IV were not tested due to absence or inadequate numbers of informative sites.

In agreement with the Neighbor-net, high compatibility scores (1.00) were observed within Enteritidis and Typhimurium (Table 3.6), indicating that most diversity within these groups might have generated by point mutations rather than recombination. Although parallel paths were observed in Paratyphi B-III by Neighbor-net, a high overall compatibility (0.98) was observed between the informative sites indicating an absence of recombination or repeated mutations. Similarly, a limited number of incompatibilities were observed in Newport-III as suggested by a compatibility score of 0.82. In contrast, the compatibility score (0.52) was much lower in Newport-II, suggesting that informative sites have different evolutionary histories in Newport-II than Newport-III. Alternatively, the high compatibility scores may reflect the low numbers of informative sites in all the groups except for Newport-II.

Table 3.6. Compatibility scores and Φw probability within each ClonalFrame group

Group	Compatibility	Φw
Enteritidis	1.00	1.00
Kentucky-I	n.d.	n.d.
Kentucky-II	n.d.	n.d.
Newport-I	n.d.	n.d.
Newport-II	0.57	<0.001*
Newport-III	0.82	<0.001*
ParatyphiB-I	n.d.	n.d.
ParatyphiB-II	n.d.	n.d.
ParatyphiB-III	0.98	0.98
ParatyphiB-IV	n.d.	n.d.
Typhimurium	1.00	1.00

n.d., not done; *, significant.

Since a high compatibility score was observed even for the groups with multiple parallel paths in Neighbor-net, another statistical program, Φw test was used. The Φw test has been compared with many other recombination detection programs and found to be the best that can also distinguish recurrent mutation from recombination (Bruen *et al.*, 2006). However, similar to the Reticulate results, recombination was not detected in Enteritidis, Paratyphi B-III and Typhimurium (Table 3.6). Recombination was only detected in Newport-II and Newport-III ($p < 0.05$).

Both the tests suggest that mutation has been the main source of diversity within Enteritidis, Paratyphi B-III and Typhimurium. Recombination introduced limited incompatibilities within Newport-III and even more incompatibilities in Newport-II. However, the power of these tests might be limited due to a very low nucleotide diversity and limited numbers of parsimony informative sites within each group. Pairwise compatibilities could not be calculated between gene fragments due to complete lack or

insufficient numbers of informative sites in many gene fragments within the groups. Φw test has also been found to have difficulties detecting recombination from the alignments where recombination rates were very low or nucleotide diversity was ≤ 1% (Bruen *et al.*, 2006).

As an alternative, the classical approach (Feil *et al.*, 2000) was used to assess the relative roles of recombination and mutation in the evolution of each group. Based on the number of nucleotide differences between alleles in related STs, the mutational and recombinational events were scored at each step of the MS_{TREE} starting from the predicted founder of each group. The ST with the highest numbers of single locus variants (SLVs) was considered to represent the founder of each group. In cases of more than one potential founder with equal numbers of SLVs, the ST with the highest number of isolates was chosen. If alleles in STs linked to the founder differed from the founder alleles by three or more nucleotides, they were classified as recombinational events. If they differed by one or two nucleotides, they were considered to result from mutations unless they were also present in other clonal groups or STs of other serovars, in which case they were scored as recombinational (Feil *et al.*, 1999; Feil *et al.*, 2000; Feil *et al.*, 2001). The same criteria were then sequentially applied to the descendents of each SLV. The ratio of recombinational to mutational events (per allele R/M) was calculated for each group. The number of polymorphic sites introduced by recombination was divided by number of nucleotides changed by mutation in each group to calculate a relative likelihood that a nucleotide will change by recombination than mutation (per site R/M).

The clonal groups identified using the BURST algorithm have been extensively used to score mutation and recombinational events by this approach (Feil *et al.*, 2001). I tested the newer version of this algorithm, eBURST 3.0, on Newport data to generate a radial diagram of SLVs linked to predicted founders (Fig. 3.4). The predicted founder represents the most likely founder genotype of the group (blue circles in Fig. 3.4). This algorithm also identifies co-founder or subgroup founder that is an SLV of the predicted founder which has diversified to generate a number of its own SLVs (yellow circles in Fig. 3.4). The structure of Newport groups were consistent with the MS_{TREE} when an eBURST group was defined as a group of three or more STs where isolates shared at least six alleles with at least one other isolate in the group. However, the relatedness between the genotypes was not completely consistent between the eBURST diagram and the MS_{TREE}. Newport-II was divided into two groups by eBURST, designated as Newport-IIa and Newport-IIb. In addition, ST158, ST184, ST193 and ST201 which are DLVs of ST45 in Newport-IIb were

separated from this subgroup. I also noted that ST193 is a DLV of ST31 in Newport-IIa that links both Newport-IIa and IIb in the MS_{TREE}. Similarly, ST123 (a DLV of ST115) and ST350 (triple locus variant, TLV, of ST118) were also separated from Newport-III by eBURST.

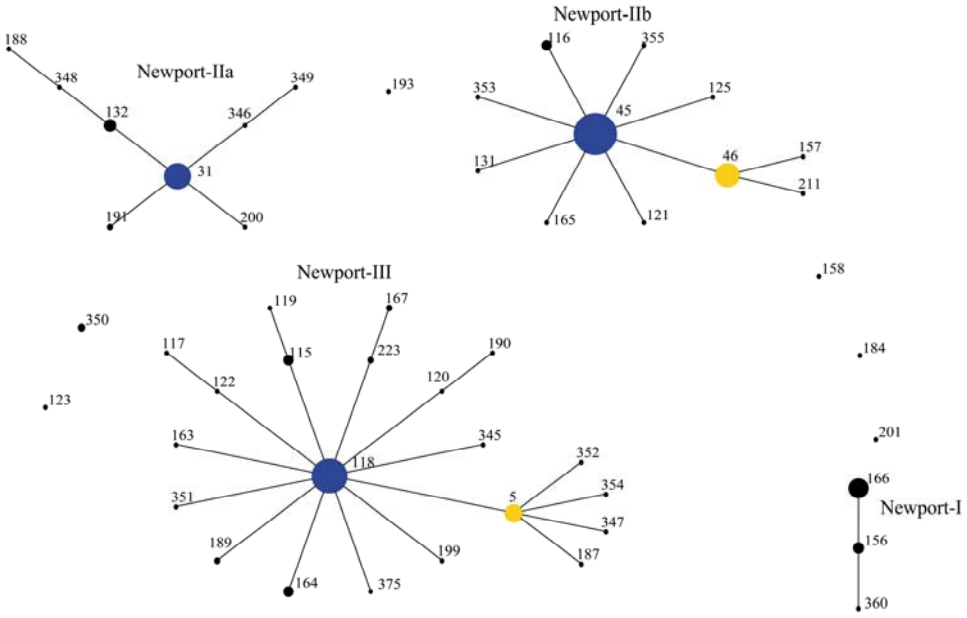

Figure 3.4. eBURST diagram on Newport isolates. Predicted founders are positioned centrally in each cluster and are shown in blue color. The subgroup founders (co-founders) are shown in yellow color. The area of each circle corresponds to the abundance of the isolates of the ST.

An MS_{TREE} of Newport isolates (Fig. 3.5) provides more information on the inter-genotype relationships because DLVs and TLVs are linked to the respective groups in addition to SLVs. Estimates of mutations and recombinational events would be more precise if these DLVs and TLVs were included. Therefore, I used the MS_{TREE} to score events.

A graphical summary of the assignments of events to mutation or recombination within each Newport group is shown in Fig. 3.5. Newport-I consists of a founder ST (ST156) plus two SLVs. The variant alleles in each of the SLVs differed from the founder alleles by one SNP each, but were scored as recombinational events because the same alleles were observed in other Newport groups and/or other serovars (Fig 3.5). The R/M ratios were > 2.0 per allele as well as per site (Table 3.7).

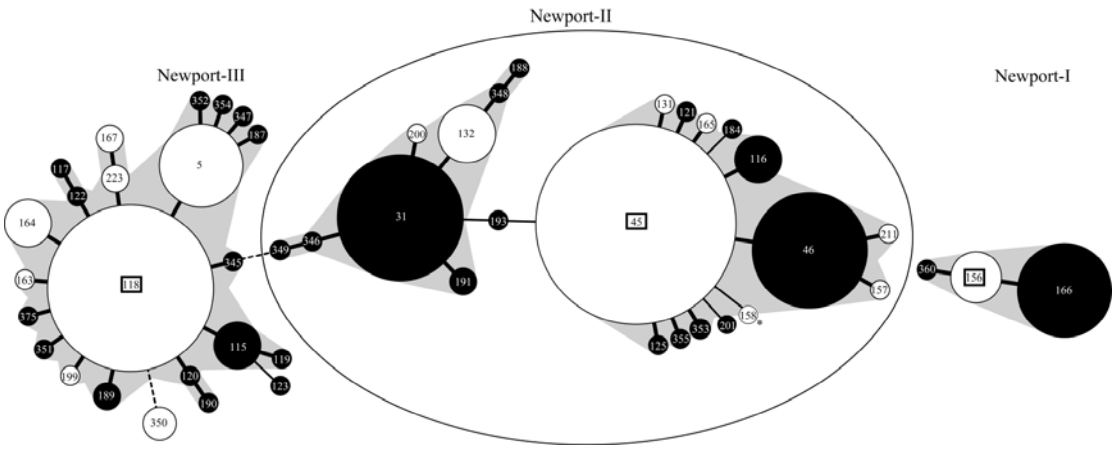

Figure 3.5. An MS_{TREE} of Newport isolates. STs in rectangles are the predicted founders. STs derived from the putative ancestors by mutational events in variant alleles are shown in white and those by recombination in black. The sizes of circles correspond to the number of isolates. In ST158 (marked by a * sign) of Newport-II, one variant allele was derived by a point mutation in founder alleles while a second arose by recombination.

Table 3.7. Recombinational and mutational events

Groups	Founder ST	Mutational events	Recombination events	R/M per allele	Polymorphic sites	R/M per site
Enteritidis	ST11	10	4	0.4	26	1.6
Kentucky-II	ST152	1	2	2.0	15	14.0
Newport-I	ST156	0	2	>2.0	2	>2.0
Newport-II	ST45	7	20	2.9	64	8.1
Newport-III	ST118	9	16	1.8	46	4.1
Paratyphi B-II	ST42	1	4	4.0	21	5.0
Paratyphi B-III	ST86	2	7	3.5	28	13.0
Paratyphi B-IV	ST88	1	2	2.0	35	34.0
Typhimurium	ST19	16	5	0.3	29	0.7

Note: Groups Kentucky-I and Paratyphi B-I consist of only single ST each, therefore, excluded.

ST45 is the founder of Newport-II, which had eight SLVs and four DLVs. The variant alleles in two of the SLVs differed by a single SNP from the founder alleles and were not found elsewhere. They were, therefore, scored as point mutations. The variant alleles in the remaining SLVs were found in other Newport groups and/or other serovars. These were scored as recombinational events. Both variant alleles in three DLVs (STs 20, 184 and 193) were present elsewhere (recombination) whereas one mutational and one recombinational event were scored for the fourth DLV (ST158).

In addition to the SLVs and DLVs directly descended from ST45, several STs were linked to the founder through an SLV (ST46) or a DLV (ST193). ST46 was the putative ancestor of STs 157 and 211 and the events were scored in these STs based on the differences from ST46. ST193 was a DLV to both ST45 as well as to ST31. ST193 was used as the ancestral genotype to score events in ST31 and ST31 was used to score events in its SLVs and

further links radiating from ST31. For Newport-II, the average per allele R/M was 2.9 and the per site R/M was 8.1. Of the 64 polymorphic sites that were observed among variant alleles, only seven were introduced by point mutations.

The founder ST (ST118) of Newport-III had 12 SLVs which were used to score mutational and recombinational events, as described above. The events were also scored in ST350 although it is a TLV to ST118; all three variant alleles were derived by point mutations. A total of 9 mutational and 16 recombinational events were scored in Newport-III leading to a per allele R/M of 1.8 and per site R/M of 4.1. Both values are slightly lower than for Newport-II.

These analyses indicate that recombination has been more important than mutation for introducing polymorphic sites in new alleles within Newport groups. Similar observations were made for Kentucky-II, Paratyphi B-II, Paratyphi B-III and Paratyphi B-IV. The values of per allele R/M were 2.0 in Kentucky-II and Paratyphi B-IV, which are comparable to Newport-III. The per allele R/M were ~ 4.0 in Paratyphi B-II and Paratyphi B-III. In contrast, the per allele R/M in Enteritidis and Typhimurium were < 1.0, indicating that most new alleles arose from mutations.

The R/M per site values indicate that an individual polymorphism is much more likely to arise by recombination than mutation in all groups except for Enteritidis and Typhimurium. The R/M per site values for the Newport groups ranged from > 2.0 to 8.1, which are higher than that of Enteritidis or Typhimurium but lower than Kentucky-II, Paratyphi B-III and Paratyphi B-IV. These results contrast to the Φw test, which only detected recombination within Newport-II and Newport-III. Neighbor-net, compatibility scores and the Φw test, all confirmed that mutations were more frequent than recombination in Enteritidis and Typhimurium.

3.3.4 Properties of Newport groups

3.3.4.1 Association with hosts

Newport lineages that seemed to be associated with humans and animals have been identified by two studies (Beltran *et al.*, 1988; Alcaine *et al.*, 2006). 105 Newport isolates were grouped into two distinct clusters by MLEE analyses and the distribution of isolates from humans and domesticated animals was significantly different between them (Beltran *et al.*, 1988). 75% of the isolates in one group were isolated from humans whereas 63% of isolates in the other group were from swine or other mammals. Similarly, two lineages were

identified when 43 Newport isolates (18 human and 25 bovine isolates) were analyzed by a three gene MLST scheme (Alcaine *et al.*, 2006). All bovine isolates clustered together in a group with 61% of human isolates. All isolates in the second cluster were isolated from humans. The distribution of isolates from both the sources in the two groups was significantly different indicating an association between these groups and the host categories. I therefore tested whether the proportions of isolates from humans and non-human hosts (avian, bovine, equine, swine and reptiles) were significantly different between the Newport groups by a χ^2 test. Isolates from food, animal feed or water were not included because these sources represent vectors rather than hosts. I also excluded isolates lacking the host information.

None of the three Newport groups that were identified in this study was completely specific to a single host. However, all 28 reptile isolates were exclusively confined to Newport-II and Newport-III (Table 3.8, Fig. 3.6A). The distribution of isolates from humans and non-human hosts into the three Newport groups differed significantly ($p = 0.01$). This might reflect the existence of Newport-I, of which 27/29 isolates were from humans. The proportion of NhWBA isolates was higher in Newport-II (83%) than in the two other groups (Table 3.8). Only 15% of NhWBA and 36% of reptile isolates grouped in Newport-III.

Table 3.8. Distribution of isolates from various hosts among Newport groups

Groups	Avian Isolates (%)	Human Isolates (%)	NhWBA* Isolates (%)	Reptiles Isolates (%)
Newport-I	1 (5)	27 (16)	1 (2)	0 (0)
Newport-II	11 (50)	44 (26)	34 (83)	18 (64)
Newport-III	10 (45)	97 (58)	6 (15)	10 (36)

* Non-human warm blooded animals, includes isolates from bovine, equine and swine hosts.
Note: 25 isolates that were isolated from food, feed, fertilizer, or unknown hosts were excluded.

3.3.4.2 Association with time

All Newport isolates were isolated between 1940 and 2005, except for one that was isolated in 1918 from a human in France. I grouped them into three categories, i) those isolated prior to 1959, ii) between 1960 and 1979 and iii) since 1980 (Table 3.9). A χ^2 test was computed on these numbers to test if the distribution of isolates among clonal groups from different time categories was significantly different.

All the isolates in Newport-I were isolated during or after the 1970s (Fig. 3.6B). However, the number of Newport-I isolates was low and no significant association was revealed between the Newport groups and time of isolation (p = 0.15).

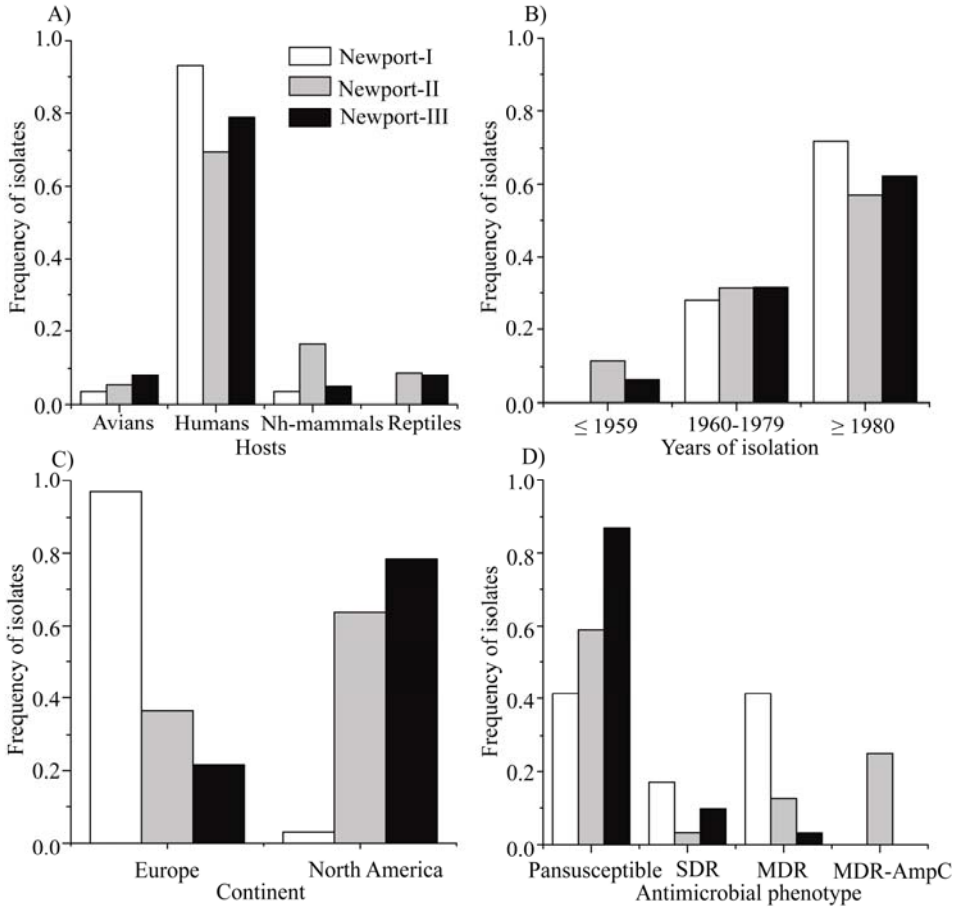

Fig. 3.6. Frequency of isolates among Newport groups from different A) hosts (excluding 25 isolates from food, feed or other unknown sources), B) dates (excluding 8 isolates with no information), C) continents (excluding 27 isolates from other geographic locations or no information) and D) antimicrobial resistance phenotype (excluding 20 isolates that were not tested).

Table 3.9. Distribution of isolates from different time of isolation among Newport groups

Groups	Number of isolates (%)		
	≤ 1959	1960-1979	≥ 1980
Newport-I	0 (0)	9 (7.7)	23 (10.2)
Newport-II	25 (75.8)	68 (58.1)	124 (54.9)
Newport-III	8 (24.2)	40 (34.2)	79 (35.0)

Note: Eight isolates with unknown dates of isolation were excluded. Association between the time of isolation and the Newport groups was not found to be significant (p = 0.15).

3.3.4.3 Geographic association of Newport groups

Most of the strains were isolated either in Europe or in North America. 31 of the 32 isolates in Newport-I were isolated in Europe whereas most isolates of the other groups were found in North America. The distribution of isolates from these continents was significantly different between the groups by a χ^2 test ($p < 0.001$) (Table 3.10, Fig. 3.6C).

Table 3.10. Distribution of isolates from Europe and North America among Newport groups

Continent	Newport-I	Newport-II	Newport-III
Europe	31 (96.9)	73 (36.7)	28 (22.2)
North America	1 (3.1)	126 (63.3)	98 (77.8)

Note: 27 isolates that were isolated in Asia, Africa, Caribbean, South America or did not have this information were excluded. Distribution of isolates from Europe and North America is significantly different among Newport groups ($p < 0$ .0001).

3.3.4.4 Antimicrobial susceptibility patterns and association with Newport groups

A total of 364 Newport isolates were tested against a panel of 16 antimicrobials. 244 isolates (67%) were pan-susceptible and the remaining isolates were resistant to at least one antimicrobial. The most prevalent resistances to individual antimicrobials among these 120 isolates were to sulfamethoxazole (28%), tetracycline (26%), streptomycin (25%) and ampicillin (24%) followed by chloramphenicol and cephalothin (19% each), amoxicillin-clavulanic acid and ceftiofur (16% each), cefoxitin (15%), kanamycin (8%), trimethoprim-sulfamethoxazole (7%), ceftriaxone (5%), gentamicin (4%) and nalidixic acid (2%) (Table 3.11).

A total of 51 antimicrobial susceptibility patterns were observed among Newport isolates (Table 3.12). These patterns were simplified into phenotype categories, e.g. multi-drug resistant (MDR), single drug resistant (SDR) and pan-susceptible. A strain was designated MDR if it was resistant to at least two separate classes of antimicrobials (Weill *et al.*, 2006) and isolates that were resistant to any single antimicrobial were designated as single drug resistant (SDR). Isolates that were resistant to extended spectrum cephalosporins (MDR-AmpC) were further separated from MDR isolates. These isolates were resistant to ampicillin, chloramphenicol, streptomycin, sulphamethoxazole, tetracycline, amoxicillin/clavulanic acid, cephalothin, cefoxitin and ceftiofur and intermediate or full resistant to ceftriaxone. However, resistance to additional antimicrobials was also observed for many MDR-AmpC isolates (Table 3.12).

96 (26%) isolates were multi-drug resistant (MDR), 53 (55%) of which were characterized as MDR-AmpC. One human isolate that was resistant to streptomycin and intermediately resistant to ampicillin, chloramphenicol and tetracycline was also designated as SDR. Another isolate from an avian host was resistant to ampicillin and cephalothin. This isolate was not assigned to any category because it was the sole isolate that was resistant to two antimicrobials which belonged to the same class. The distribution of isolates with these phenotypes was tested for any association with human or non-human sources by a χ^2 test. Isolates from food, animal feed, fertilizer and meat were excluded for the reasons previously mentioned.

Table 3.11. Antimicrobial resistance among Newport isolates from different host types

Antimicrobial agent	Resistance Breakpoint (μg/ml)	Resistant isolates (intermediate resistant)							
		Human (258)	Bovine (20)	Swine (16)	Chicken (13)	Turkey (9)	Reptile (27)	Others[a] (21)	Total (364)
Ampicillin	≥ 32	57 (1)	16 (0)	7 (0)	4 (0)	1 (0)	0 (0)	4 (0)	89 (1)
Amoxicillin/ clavulanic acid	≥ 32/16	27 (2)	16 (0)	7 (0)	3 (0)	0 (0)	0 (0)	4 (0)	57 (2)
Ceftriaxone	≥ 64	16 (12)	1 (15)	0 (7)	0 (3)	0 (0)	0 (0)	0 (4)	17 (41)
Cephalothin	≥ 32	37 (1)	16 (0)	7 (0)	4 (0)	0 (0)	0 (0)	4 (0)	68 (1)
Chloramphenicol	≥ 32	39 (1)	17 (0)	6 (0)	4 (0)	0 (0)	0 (0)	4 (0)	70 (1)
Trimethoprim/ sulfamethoxazole	≥ 4/76	17 (0)	3 (0)	2 (0)	0 (0)	0 (0)	0 (0)	3 (0)	25 (0)
Cefoxitin	≥ 32	26 (1)	16 (0)	7 (0)	3 (0)	0 (0)	0 (0)	4 (0)	56 (1)
Gentamicin	≥ 16	6 (0)	4 (0)	2(0)	1 (0)	2 (0)	0 (0)	0 (0)	15 (0)
Kanamycin	≥ 64	14 (0)	8 (0)	3 (0)	2 (0)	1 (0)	0 (0)	0 (0)	28 (0)
Nalidixic acid	≥ 32	7 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (0)	8 (0)
Sulfamethoxazole	≥ 512	64 (0)	18 (0)	8 (0)	4 (0)	2 (0)	1 (0)	4 (0)	101 (0)
Streptomycin	≥ 64	57 (0)	18 (0)	8 (0)	4 (0)	1 (0)	0 (0)	4 (0)	92 (0)
Tetracycline	≥ 16	56 (1)	19 (0)	8 (0)	5 (0)	0 (0)	0 (0)	5 (0)	93 (1)
Ceftiofur	≥ 8	28 (0)	16 (0)	7 (0)	3 (0)	0 (0)	0 (0)	4 (0)	58 (0)

All the isolates were sensitive to amikacin and ciprofloxacin (resistance breakpoints ≥ 32 and ≥ 4 μg/ml respectively).

[a] Isolates from animal feed, food, fertilizer, frog legs, horse, lion, meat and rat.

MIC values (μg/ml) for intermediate resistance to ampicillin (16), amoxicillin/clavulanic acid (16/8), ceftriaxone (16-32), cephalothin (16), chloramphenicol (16), cefoxitin (16), gentamicin (8), kanamycin (32), streptomycin (16), tetracycline (8) and ceftiofur (4)

Table 3.12: Antimicrobial resistance patterns of Newport isolates from different sources

Resistance profile	Simple phenotype	Number of isolates (%)						
		Human	Bovine	Swine	Chicken	Turkey	Reptile	Others
AmpAugChlStrTetTioSmxFoxCepSxtGen[Axo]Kan	MDR-AmpC	0 (0)	1 (5)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
AmpAugChlStrTetTioSmxFoxCep[Axo]GenKan	MDR-AmpC	0 (0)	3 (15)	1 (6.3)	1 (7.7)	0 (0)	0 (0)	0 (0)
AmpAugChlStrTetTioSmxFoxCep[Axo]SxtKan	MDR-AmpC	0 (0)	1 (5)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
AmpAugChlStrTetTioSmxFoxCepAxoGenKan	MDR-AmpC	2 (0.8)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
AmpAugChlStrTetTioSmxFoxCep[Axo]Kan	MDR-AmpC	1 (0.4)	0 (0)	0 (0)	1 (7.7)	0 (0)	0 (0)	0 (0)
AmpAugChlStrTetTioSmxFoxCep[Axo]Sxt	MDR-AmpC	5 (1.9)	1 (5)	1 (6.3)	0 (0)	0 (0)	0 (0)	3 (14.3)
AmpAugChlStrTetTioSmxFoxCepAxoKan	MDR-AmpC	0 (0)	1 (5)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
AmpAugChlStrTetTioSmxFoxCepAxoSxt	MDR-AmpC	1 (0.4)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
AmpAugChlStrTetTioSmxFoxCepAxo	MDR-AmpC	10 (3.9)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
AmpAugChlStrTetTioSmxFoxCep[Axo]	MDR-AmpC	5 (1.9)	9 (45)	4 (25)	1 (7.7)	0 (0)	0 (0)	1 (4.8)
AmpAugStrTetTioSmxFoxCep[Axo]SxtKan	MDR	0 (0)	0 (0)	1 (6.3)	0 (0)	0 (0)	0 (0)	0 (0)
AmpAugChlStrTetTioSmx[Fox]Cep[Axo]	MDR	1 (0.4)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
AmpStrTetTioSmxFoxCepAxoKanNal	MDR	1 (0.4)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
AmpAugStrTetTioSmxFoxCepAxo	MDR	1 (0.4)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Amp[Aug]ChlStrTetSmxKan	MDR	1 (0.4)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
AmpChlStrTetSmxSxtGenNal	MDR	1 (0.4)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
AmpStrTetSmx[Cep]GenKan	MDR	1 (0.4)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
AmpAugChlStrTetSmxKan	MDR	1 (0.4)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
AmpChlTetSmxSxtGenNal	MDR	1 (0.4)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
AmpChlStrTetSmxSxtKan	MDR	4 (1.6)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
AmpChlStrTetSmxSxtNal	MDR	1 (0.4)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
AmpChlStrTetSmxSxt	MDR	2 (0.8)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
AmpChlStrTetSmxCep	MDR	1 (0.4)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
AmpStrTetSmxCepKan	MDR	1 (0.4)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
AmpTioSmxCepAxoSxt	MDR	1 (0.4)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Amp[Aug]StrTetCep	MDR	1 (0.4)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
AmpChlStrTetSmx	MDR	1 (0.4)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
AmpStrTetSmxCep	MDR	5 (1.9)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
AmpChlTetSmxNal	MDR	1 (0.4)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)

Continued Table 3.12…

Resistance profile	Simple phenotype	Number of isolates (%)						
		Human	Bovine	Swine	Chicken	Turkey	Reptile	Others
ChlStrTetSmxKan	MDR	0 (0)	1 (5)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
StrTetSmxGenKan	MDR	0 (0)	0 (0)	1 (6.3)	0 (0)	0 (0)	0 (0)	0 (0)
AmpSmxGenKan	MDR	0 (0)	0 (0)	0 (0)	0 (0)	1 (11.1)	0 (0)	0 (0)
AmpSmxSxtKan	MDR	1 (0.4)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
AmpStrTetSmx	MDR	1 (0.4)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
AmpStrSmxCep	MDR	1 (0.4)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
ChlStrTetSmx	MDR	0 (0)	0 (0)	0 (0)	1 (7.7)	0 (0)	0 (0)	0 (0)
StrTetSmxKan	MDR	0 (0)	1 (5)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
AmpStrTet	MDR	1 (0.4)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
StrSmxGen	MDR	1 (0.4)	0 (0)	0 (0)	0 (0)	1 (11.1)	0 (0)	0 (0)
StrTetSmx	MDR	1 (0.4)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
AmpStr	MDR	1 (0.4)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
SmxKan	MDR	1 (0.4)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
StrTet	MDR	1 (0.4)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Amp	SDR	2 (0.8)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Nal	SDR	2 (0.8)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (4.8)
Smx	SDR	9 (3.5)	0 (0)	0 (0)	0 (0)	0 (0)	1 (3.7)	0 (0)
Str	SDR	2 (0.8)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Str [AmpChlTet]	SDR	1 (0.4)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Tet	SDR	3 (1.2)	1 (5)	0 (0)	1 (7.7)	0 (0)	0 (0)	1 (4.8)
AmpCep	n.a.	0 (0)	0 (0)	0 (0)	1 (7.7)	0 (0)	0 (0)	0 (0)
Pan-susceptible	Pan	180 (69.8)	1 (5)	8 (50)	7 (53.8)	7 (77.8)	26 (96.3)	15 (71.4)

Note: 1. An human isolate was resistant to streptomycin and intermediate resistant to ampicillin, chloramphenicol and tetracycline was also designated as SDR.
2. The abbreviation of antimicrobials in parentheses in the cloumn “Resistance profile” indicates the intermediate resistance of the strain to that antimicrobial.
n.a., not assigned. Because it was the only isolate that was resistant to two antimicrobials which belonged to the same class.
Abbreviations: Amp, ampicillin; Aug, amoxicillin/clavulanic acid; Axo, ceftriaxone; Cep, cephalothin; Chl, chloramphenicol; Sxt, trimethoprim/sulfamethoxazole; Fox, cefoxitin; Gen, gentamicin; Kan, kanamycin; Nal, nalidixic acid; Smx, sulfamethoxazole; Str, streptomycin; Tet, tetracycline; Tio, ceftiofur.

The distribution of isolates with different antimicrobial phenotype was significantly different between human and non-human isolates (p = 0.0002). 180 human isolates (70%) were pan-susceptible and only 59 (23%) were MDR. 24 (41%) of human MDR isolates were MDR-AmpC. In contrast, 36% (32 of 89) of isolates from animal sources were MDR and 78% of them were MDR-AmpC. However, most data for animal isolates in this study was from Harbottle *et al*., 2006 who primarily collected isolates from clinically ill food animals in the U.S.A. MDR-AmpC has been a major problem among food animals in the United States (Zhao *et al.*, 2001; Devasia *et al.*, 2005). I analyzed only 36 additional isolates from animal sources, 27 of which were reptile in origin. Most reptile isolates (96%) were pan-susceptible.I also tested whether the proportions of isolates with MDR (including MDR-AmpC isolates), SDR and pan-susceptible phenotypes were significantly different between the Newport groups by χ^2 test (Table 3.13).

Table 3.13. Distribution of isolates of different resistance phenotypes among Newport groups

Groups	Number of isolates (%)			
	MDR-AmpC	MDR	SDR	Pan-susceptible
Newport-I	0 (0)	12 (27.9)	5 (21.7)	12 (4.9)
Newport-II	53 (100.0)	27 (62.8)	7 (30.4)	125 (51.2)
Newport-III	0 (0)	4 (9.3)	11 (47.8)	107 (43.9)

Note: 20 isolates that were not tested for antimicrobials were not included

The distribution of isolates with these phenotypes was significantly different among the Newport groups (p < 0.0001). All MDR-AmpC isolates were in Newport-II and the highest proportion of other MDR isolates (63%) was also in Newport-II. Within Newport-II, the MDR-AmpC phenotype was restricted to STs 45 and 116 (Fig. 3.6D) which are SLVs. In contrast, most of the isolates in Newport-III were pan-susceptible and Newport-I contained equal numbers of MDR and pan-susceptible isolates.

3.4 Discussion

3.4.1 Newport, a polyphyletic serovar

Previous studies have indicated that Newport is a polyphyletic serovar (Beltran *et al.*, 1988; Sukhnanand *et al.*, 2005; Torpdahl *et al.*, 2005; Harbottle *et al.*, 2006). But these studies involved smaller numbers of isolates and less powerful methods. An MLEE analysis of 105 isolates grouped them into two distinct clusters, one associated with humans and the second with the domesticated animals (Beltran *et al.*, 1988). Similarly, two lineages were identified

when 43 Newport isolates were analyzed by a three gene MLST scheme (Alcaine *et al.*, 2006).

In this study, I defined three lineages, Newport-I, Newport-II and Newport-III, after analyzing 384 Newport isolates by a seven gene MLST scheme. Three different approaches were used on the MLST data to infer population structure of this serovar. First was ClonalFrame (Didelot and Falush, 2006) which deduces genealogies from multilocus sequence data with only minimal effects of recombination by identifying and excluding potentially imported stretches. Three distinct groups of Newport were identified by this program (Fig. 3.1). An algorithm called Neighbor-net (Bryant and Moulton, 2004) which accounts for recombination and resolve conflicting signals as parallel paths in a phylogenetic network was also applied to the multilocus sequence data. The best-fit model of nucleotide substitution was used in order to infer the most accurate evolutionary relatedness between the STs. Neighbor-net also identified three groups within Newport that were similar to the ClonalFrame groups in composition (Fig. 3.2). However, Neighbor-net separated three STs from Newport-II, possibly due to alleles that were imported from other serovars.

The MS_{TREE} application of Bionumerics 4.5 (Applied Maths, Sint-Martens-Latem, Belgium) was also applied to the data. It identifies groups of closely related genotypes based on the allelic profile of isolates and links STs at unique minimal distances. This application uses some of the principles of the BURST algorithm which has been widely used in bacterial population genetics to study evolutionary relationship using the MLST data (Enright *et al.*, 2002; Feil *et al.*, 2004; Honsa *et al.*, 2008; Feil *et al.*, 2003). The MS_{TREE} also identified three groups of identical composition to ClonalFrame (Fig. 3.3). Although Newport-II was split into two subgroups and singletons, they were closely related when groups were defined to contain ≥ 3 STs where an ST shared a minimum of five alleles with at least one other ST of the group (Fig. 3.3). A close association between the STs of both the subgroups was also indicated by a thick net of cross-links. Two STs were separated from Newport-III in the MS_{TREE} but their cross-linkage was to this group and they did not possess strong similarity to genotypes in other groups.

Therefore, all the approaches indicated the existence of three distinct lineages within serovar Newport. However, multiple STs of Newport-II and Newport-III were connected by cross-links in the MS_{TREE}. This suggests recombination but it is not clear whether these groups are merging through homologous recombination such as *Campylobacter jejuni* and

C. coli (Dingle *et al.*, 2005) or drifting apart by the acquisition of additional diversity from other serovars.

In the MLEE study, SARB37 (Np11) and SARB38 (Np15) belonged to one lineage while SARB36 (Np8) belonged to a second lineage (Beltran *et al.*, 1988). In this study, these belonged to STs 31, 46 and 5, respectively. STs 31 and 46 belong to Newport-II and ST5 to Newport-III. Therefore, the two lineages identified by the MLEE study probably correspond to Newport-II and Newport-III.

A third lineage, called Newport-I, is reported in this study but has not been previously described. None of the alleles in this lineage were common to the other two Newport lineages except that *hisD*12 and *thrA*12 were found in several STs of Newport-II as well as some STs from other serovars. This lineage was distinct by all three approaches that were used to infer evolutionary relatedness. None of the STs of this lineage was cross-linked to any ST of Newport-II, Newport-III or the groups of other serovars. Newport-I comprised of only three STs that differed from each other by one or two polymorphic nucleotide sites (SNPs). Nucleotide diversity in this lineage was also significantly lower than other two Newport groups. The variant alleles among these STs were assumed to have originated by recombinational events because they were found in other serovars. However, these alleles differed from the founder alleles only by one nucleotide and may represent homoplasies associated with independent point mutations. These results suggest that Newport-I represents a discrete lineage which has recently arisen by lateral transfer of genes encoding antigenic determinants (Beltran *et al.*, 1988; Selander *et al.*, 1994). Alternatively, the strains of this group are derived from a recent evolutionary bottleneck which removed most nucleotide diversity from this lineage.

In contrast to these population genetic analyse based on sequence homology, a study based on microarray hybridization experiments proposed that Newport is monophyletic (Porwollik *et al.*, 2004). A total of five Newport isolates (including SARB36, SARB37 and SARB38) and 74 isolates of other serovars were hybridized to PCR amplified sequences covering 94.5% of the Typhi genome and 96.6% of the Typhimurium LT2 genome. The clustering behaviour of most serovars was similar to that of MLEE data except for Newport and Muenchen (Porwollik *et al.*, 2004). Newport isolates were monophyletic and shared a clade with isolates of serovar Muenchen. Both these serovars were polyphyletic based on MLEE analyses and no relationship was observed between them (Beltran *et al.*, 1988; Selander *et al.*, 1990b). Similarly, no correlation was observed between isolates of both serovars based on MLST analysis (unpublished data). At least five alleles were different

between the Newport and the Muenchen STs (data available at the *Salmonella* MLST website). The extent of sequence diversity within a serovar may not be reflected by simply scoring genes for presence or absence. Furthermore, phylogenetic relationship among strains based on the presence or absence of gene clusters may become blurred when genes are acquired by recombination between serovars (Porwollik *et al.*, 2004). SARB36, SARB37 and SARB38 were separated into two lineages (Newport-II and Newport-III) on the basis of MLST in this study. I did not perform MLST with the particular strains that were used in the microarray experiments. Since missassignments and strain swapping of many strains of the SARB collection have been reported (Porwollik *et al.*, 2004), those particular SARB strains should be tested by MLST in order to compare microarray hybridization results with MLST analyses.

3.4.2 Association of Newport groups with hosts, time, geography and antimicrobial resistance phenotype

The three Newport lineages were found to be differentially associated with human versus non-human hosts (avian, bovine, equine, swine and reptiles) in this study ($p = 0.01$). Most isolates of Newport-I were from humans. A higher proportion of isolates from non-human sources was found in Newport-II than in Newport-III whereas the proportion of human isolates was higher in Newport-III.

Two Newport lineages have been reported that were associated with human and domesticated animals, respectively (Beltran *et al.*, 1988; Alcaine *et al.*, 2006). According to MLEE, 105 Newport isolates from humans and domesticated animals (swine or other mammals) were in two lineages that differed significantly in the distribution of isolates from these sources (Beltran *et al.*, 1988). Similarly, according to MLST, 43 Newport isolates from human and bovine hosts were in two distinct clusters and the distribution of isolates from human and bovine sources was significantly different between the clusters (Alcaine *et al.*, 2006). Consistent with these observations, in this study, most bovine (90%) and swine (87.5%) isolates were in group Newport-II. Furthermore, the proportions of human isolates were higher in two lineages, Newport-I and Newport-III than in Newport-II. Although isolates in Newport-I have only been isolated since the 1970s, no significant association existed between the time of isolation and the phylogenetic grouping ($p = 0.15$). Possibly, the distribution of isolates in the three Newport groups was independent of the time of isolation. Alternatively, the lack of significance reflects the low number of Newport-I isolates.

The distribution of isolates from Europe and North America varied between the Newport groups ($p < 0.001$). Newport-I is particularly prevalent in Europe. The proportions of isolates from North America were higher than from Europe for both Newport-II and Newport-III. Furthermore, the relative frequency of isolates from Europe was more than 2fold higher in Newport-II (55%) than in Newport-III (21%). This observation suggests that Newport-III is more common in the U.S.A. than in Europe.

Antimicrobial susceptibility phenotype was associated with host ($p = 0.0002$). The proportion of pan-susceptible isolates was higher among humans whereas the MDR-AmpC phenotype was more common among animal isolates. However, this observation is partially due to a high proportion of animals isolates that were isolated from food animals in the U.S.A. where infection of MDR-AmpC Newport has been a major problem among food animals (Zhao *et al.*, 2001; Devasia *et al.*, 2005).

The proportion of MDR and pan-susceptible isolates was significantly different between the Newport groups ($p < 0.0001$). While 38% of Newport-II isolates were MDR, most Newport-III isolates were pan-susceptible. Furthermore, 83% of MDR isolates were in Newport-II. The MDR-AmpC phenotype was restricted to STs 45 and 116 of Newport-II in agreement with previous studies suggesting a single global recent origin of MDR-AmpC strains (Harbottle *et al.*, 2006; Alcaine *et al.*, 2005; Egorova *et al.*, 2008).

3.4.3 Newport versus other serovars

Serovars of subspecies *enterica* vary in population structure (Beltran *et al.*, 1988). Based on MLEE studies, Typhimurium is monophyletic whereas serovars Enteritidis and Paratyphi B are polyphyletic (Beltran *et al.*, 1988; Selander *et al.*, 1990a; Selander *et al.*, 1990b). The 17 ETs found within 299 Typhimurium isolates were grouped in a single cluster (Beltran *et al.*, 1988). In contrast, Enteritidis was concluded to be polyphyletic by MLEE because only 10 ETs clustered together and 4 others fell into three distantly related groups (Beltran *et al.*, 1988). Similarly, 14 ETs of serovar Paratyphi B fell into three groups (Selander *et al.*, 1990a; Selander *et al.*, 1990b). Typhimurium was also concluded to be monophyletic according to MLST studies (Sukhnanand *et al.*, 2005; Torpdahl *et al.*, 2005). However, only 11 Typhimurium isolates had been analyzed in these prior MLST studies. Five isolates of serovar Kentucky were separated into two phylogenetically distant STs in one MLST study (Sukhnanand *et al.*, 2005).

The results of my study are in general agreement with the MLEE and prior MLST results. However, the overall structure of Typhimurium was comparable to that of Enteritidis. 13 STs of serovar Enteritidis formed a single cluster and only two other STs, each with one isolate, were genetically distinct (singletons) (Table 3.2). Similarly, most STs of serovar Typhimurium clustered together, except for two STs that were identified as singletons. Two groups were observed within serovar Kentucky, consistent with the previous observation (Sukhnanand *et al.*, 2005). STs of Paratyphi B isolates formed four distinct groups one of which was a single ST containing 16 isolates. Newport differed from all these serovars because the 49 Newport STs fell into three distinct groups. Therefore, the population structure of subspecies *enterica* serovars varies with serovar. The degree of polyphyletism within serovar Newport was higher than Enteritidis, Typhimurium and Kentucky but lower than Paratyphi B.

The degree of polymorphism among gene fragments also varied greatly between the serovars as well as between groups within each serovar (Table 3.5a and 3.5b). Various studies suggested that recombination has been the major source of variation among housekeeping genes within subspecies *enterica* (Brown *et al.*, 2003; Octavia and Lan, 2006; Falush *et al.*, 2006). Neighbor-net also resolved conflicting signals in groups except for Enteritidis and Typhimurium, suggesting that multiple alleles in most groups have arisen by recombination. Therefore, two programs, Reticulate (Jakobsen and Easteal, 1996) and Φw test (Bruen *et al.*, 2006), were employed to test whether recombination has been the key factor behind the current diversity within these groups. Both the programs are based on compatibilities between informative sites and are considered as the powerful methods to detect recombination from nucleotide sequences (Bruen *et al.*, 2006). However, neither of these programs detected recombination in any group except for Newport-II and Newport-III. Only limited incompatibilities were observed in Newport-III by Reticulate, indicating a lower frequency of recombination than in Newport-II. A very low nucleotide diversity and the limited number of informative sites within each group might explain why only negligible incompatibilities was observed in the other groups.

An alternative allele-based approach (Feil *et al.*, 2000) was used to assess the relative role of recombination and mutation in the clonal diversification of each group. Alleles derived by mutation or recombinational events were identified based on the nucleotide differences between the variant alleles and those in their putative ancestral genotypes. This approach has previously been used to estimate mutation and recombination events within groups of closely related isolates for various pathogenic bacteria (Feil *et al.*, 1999; Feil *et al.*, 2000;

Feil *et al.*, 2001; Feil *et al.*, 2003). Although traditionally such events were only scored between putative ancestors and their descendents SLVs, I also included DLVs or TLVs within each group because these were closely related by all the methods that were used to identify genealogies.

Recombination is more likely to introduce new alleles in Newport-II than Newport-III (Table 3.7). The exact per allele R/M value could not be determined for Newport-I because both variant alleles resulted by recombination, leading to an R/M ratio of > 2.0. The relative frequencies of recombination to mutation (per allele R/M) also varied markedly among the groups and between serovars. Point mutations have been the main source of new alleles within Enteritidis and Typhimurium whereas most variant alleles were imported from elsewhere within Paratyphi B-II and Paratyphi B-III (Table 3.7). An allele in Kentucky-II, Newport-I, Newport-II, Newport-III and Paratyphi B-IV is also more likely to change by recombination but the frequencies were intermediate i.e. higher than Enteritidis and Typhimurium and lower than Paratyphi B-II and Paratyphi B-III.

The average numbers of nucleotide sites changed by recombination than mutation (per site R/M) in Enteritidis and Typhimurium correlated with their low nucleotide diversity and low R/M per allele (Table 3.7). This correlation was also observed for Newport-II and Newport-III. The values of π, per allele R/M as well as per site R/M were higher in Newport-II than for Enteritidis, Typhimurium or Newport-III. All three values were intermediate in Newport-III, i.e., higher than for Enteritidis and Typhimurium but lower than for Newport-II. However, any correlation between per site R/M and the π values or per allele R/M was not observed for the remaining groups. I could not estimate an exact per site R/M value for Newport-I, for the same reasons as for per allele R/M values.

Although nucleotide diversity in Paratyphi B-II was comparable to Newport-II, the likelihood of generating new alleles by recombination (per allele R/M) was much higher but the likelihood of a polymorphic nucleotide to be introduced by a recombinational event (per site R/M) was lower than for Newport-II. These observations suggest that a recombinational event in Paratyphi B-II introduced comparatively low number of polymorphic nucleotides than in Newport-II. Similarly, per site R/M was much higher in Kentucky-II and Paratyphi B-IV than Paratyphi B-II despite the 2fold lower per allele R/M values and significantly lower nucleotide diversities in both the groups. Therefore, a recombination event in Kentucky-II and Paratyphi B-IV introduced alleles with more polymorphic nucleotides than in Paratyphi B-II. The value of per site R/M has been used as an indicator of sequence divergence in different bacterial species, e.g. a high per site R/M

correlated with the high nucleotide diversity in *Neisseria meningitidis* whereas the lower per site R/M values correlated with the low nucleotide diversity in *Streptococcus pneumoniae* and *Staphylococcus aureus* (Feil *et al.*, 2001). The average number of polymorphic nucleotides introduced by recombination has been highly variable between the groups of subspecies *enterica* serovars. Perhaps such an approach is not as suitable for analyses within groups of closely related genotypes within a species.

In agreement with previous observations (Brown *et al.*, 2003; Octavia and Lan, 2006; Falush *et al.*, 2006), recombination was quite frequent among housekeeping genes within subspecies *enterica*. However, the relative role of recombination and mutation in the clonal diversification has been markedly variable not only between serovars but also between the groups within a serovar. Subspecies *enterica* represents a pool of diverse strains that have been assigned to > 1500 serovars based on the antigenic profiles of the surface antigens (Popoff *et al.*, 2004). However, a serovar does not necessarily indicate a group of genetically identical isolates and genes encoding surface antigens are probably exchanged quite frequently within the subspecies giving rise to multiple lineages within a serovar.

4. Comparison of carried and disease associated salmonellae

4.1 Introduction

Non-typhoidal salmonellosis generally results in a self limiting diarrhea in humans but can also lead to a life-threatening bacteremia in immunocompromised patients (Hohmann, 2001; Gordon, 2008). The enteric invasion of salmonellae triggers immune responses to limit the spread of infection (Uzzau *et al.*, 2000; Srinivasan and McSorley, 2006). The bacteria that survive the immune responses can be disseminated to systemic organs, e.g., the liver, gallbladder and spleen causing bacteremia in patients (Worley *et al.*, 2006; Ly and Casanova, 2007).

After clinical recovery from *Salmonella* infection, 1-5% of patients continue to harbour these bacteria in their hepatobiliary system and are designated as chronic carriers (D'Aoust, 1991; Gupta *et al.*, 2006). Asymptomatic human carriers of *Salmonella* shed bacteria in their stools for years and impose a potential threat to the healthy community. Healthy carriers may be the major source of *Salmonella* infection, especially carriers that are associated with food handling (Dryden *et al.*, 1994; Ollinger-Snyder and Matthews, 1996; Kariuki *et al.*, 2006). However, it has not yet been determined whether salmonellae from chronic human carriers are the same as those isolated from non-carrier hosts.

In this study, I analyzed subspecies *enterica* isolates that had been isolated from chronic human carriers in the 1980s in Germany. In order to determine the prevalence of carrier genotypes among human clinical and animal isolates, I focused on carried isolates from serovars where MLST data were also available from non-carrier humans or animals (other mammals, birds and reptiles).

4.2 Methods

4.2.1 Bacterial isolates

Forty-five strains isolated from chronic human carriers in the 1980s in Germany were obtained from Dr. Roy Curtiss, Center for Infectious Diseases and Vaccinology, Arizona State University, Tempe, U.S.A. These isolates belonged to 20 serovars of subspecies *enterica*. I also included one Newport isolate from a human carrier from chapter 3 and one carrier isolate of serovar Abony from the *Salmonella* MLST website, resulting in 47 isolates in 22 serovars. These results were compared with MLST data of the same serovars for 686 isolates from non-carrier humans and 329 isolates from animals, including the

Newport data from chapter 3 and reptile data from chapter 5. These MLST data were downloaded from the *Salmonella* MLST website (http://web.mpiib-berlin.mpg.de/mlst/dbs/Senterica). The numbers of isolates are summarized in Table 4.1.

Table 4.1. List of isolates from different hosts

Serovar; mnemonic	No. of isolates		
	Human carriers	Non-carrier humans	Animals
Abony; Ab	1	1	0
Agona;, Ag	1	10	1
Anatum; An	3	1	3
Bovismorbificans; Bv	2	2	7
Braenderup; Bu	1	4	5
Brandenburg; Ba	1	6	1
Bredeney; Bd	1	2	0
Derby; De	1	4	3
Enteritidis; En	1	123	24
Hadar; Hd	3	11	9
Heidelberg; He	2	11	4
Infantis; In	4	11	4
Manhattan; Mn	2	4	1
Montevideo; Mo	1	9	4
Newport; Np	1	267	91
Ohio; Oh	3	0	1
Panama; Pn	2	4	5
Paratyphi B[1];Pb	1	24	18
Schwarzengrund; Sc	1	2	0
Thompson; Th	1	7	3
Typhimurium; Tm	13	164	141
Virchow; Vi	1	19	4
Total	47	686	329

[1]Including Paratyphi B var. Java

The cultivation and storage of isolates, isolation of DNA, PCR amplification, sequencing and MLST analyses were carried out as previously described (Chapter 3). An MS_{TREE} was generated in Bionumerics 4.5 based on the allelic profiles of isolates.

4.2.2 Characteristics of housekeeping genes

Nucleotide diversity (π) with Jukes-Cantor correction was calculated using MEGA 4.0 (Tamura *et al.*, 2007) separately for concatenated sequence alignments of isolates from carrier humans, non-carrier humans and non-carrier humans plus animals (ncHA). The average pairwise distance at non-synonymous (Ka) and synonymous (Ks) sites was calculated for alignments of gene fragments for isolates from human carriers and ncHA using DnaSP 4.0 (Rozas *et al.*, 2003).

4.2.3 Statistical analyses

To compare the genotypic diversity within salmonellae from human carriers, the expected numbers of STs were estimated for 47 isolates from non-carrier humans as well as from ncHA by rarefaction analysis using PAST 1.73 (Hammer *et al.*, 2001). Fisher's exact tests and χ^2 tests were computed using the statistical package STATISTICA (StatSoft, Inc., Tulsa, U.S.A.).

4.3 Results

4.3.1 Properties of housekeeping genes

The genetic diversity (π) of concatenated sequence alignments of MLST housekeeping genes was compared between carried isolates and isolates from non-carrier humans as well as ncHA. The π value for carried isolates was not significantly different from the value for ncHA isolates (Table 4.2). It also did not differ from isolates from non-carrier humans. Therefore, there does not seem to be any difference in overall selection pressure on housekeeping gene diversity between salmonellae from human carriers and other hosts.

Table 4.2. Diversity among isolates in different host categories

Host	No. of Isolates	No. of STs	π	Rarefied STs[1]
Carrier humans	47	23	0.011 ± 0.001	23
Non-carrier humans	686	101	0.010 ± 0.001	22.1 ± 2.7
ncHA	1015	137	0.009 ± 0.001	22.7 ± 2.7

[1]Number of STs estimated by rarefaction analysis for 47 isolates

The ratio of non-synonymous substitutions per non-synonymous site to synonymous substitutions per synonymous site (Ka/Ks) was also calculated on alignments of individual gene fragments to test whether any of these housekeeping genes experienced different selection pressures. The Ka/Ks value was < 1.0 for all the gene fragments in both carried and ncHA isolates, indicating the presence of purifying selection (Table 4.3).

The Ka/Ks was much higher for *hemD* than for other gene fragments (as described in chapter 3). This observation indicates that non-synonymous substitutions are fixed at higher frequency in *hemD* than in other gene fragments, although not frequently enough to suggest diversifying selection.

Table 4.3. Average pairwise distances at non-synonymous and synonymous sites among carried and ncHA isolates

	Carrier Isolates			ncHA isolates		
Genes	Ka	Ks	Ka/Ks	Ka	Ks	Ka/Ks
aroC	0.00077	0.02808	0.027	0.00010	0.02318	0.004
dnaN	0.00122	0.03543	0.034	0.00137	0.02256	0.061
hemD	0.00540	0.02415	0.224	0.00518	0.02120	0.244
hisD	0.00242	0.05839	0.041	0.00209	0.04613	0.045
purE	0.00057	0.04149	0.014	0.00083	0.03197	0.026
sucA	< 0.00011	0.03241	< 0.003	0.00009	0.03042	0.003
thrA	< 0.00011	0.05603	< 0.002	0.00004	0.06551	0.001

4.3.2 Diversity between carried and ncHA isolates

The 47 isolates from chronic human carriers fell into 23 STs. All carried isolates within a serovar possessed an identical MLST profile, except for serovar Hadar where two STs were identified among three carried isolates. In contrast, multiple STs were identified for isolates from non-carrier humans and animals for all serovars except Abony, Anatum, Heidelberg, Panama, Schwarzengrund and Thompson.

Rarefaction curves are often used to compare the number of taxa between samples of different sizes because this approach allows estimating the number of taxa that would be expected for a smaller sample from the data for larger populations (Gotelli and Colwell, 2001). Rarefaction curves were computed on ncHA isolates and also separately for isolates from non-carrier humans to predict diversity in samples of the same size as that of the isolates from human carriers. The predictions for 47 isolates were similar to those observed for isolates from carrier humans, both for non-carrier humans as well as ncHA, indicating similar levels of genotype richness between these classes (Table 4.2, Fig. 4.1).

4.3.3 Prevalence of carrier-STs among ncHA isolates

An MS_{TREE} revealed that the STs present among carried isolates were generally the most common ST for that serovar among isolates from non-carrier humans and animals (Fig 4.2). STs of all ncHA isolates were identical to carried isolates within serovars Abony, Anatum, Heidelberg, Panama, Schwarzengrund and Thompson (Table 4.4a). However, the proportion of ncHA isolates in carrier STs varied from 0.06-0.91 for the remaining serovars (Table 4.4b). For two serovars, the proportions can not be readily evaluated because only few ncHA isolates were available. Only two ncHA isolates were available for serovar

Bredeney, both from non-carrier humans. One of these ncHA isolates possessed the same ST as the sole carried isolate while the second possessed a different ST. For serovar Ohio, the sole animal isolate available was of a different ST than the three carried isolates. These are not discussed further below.

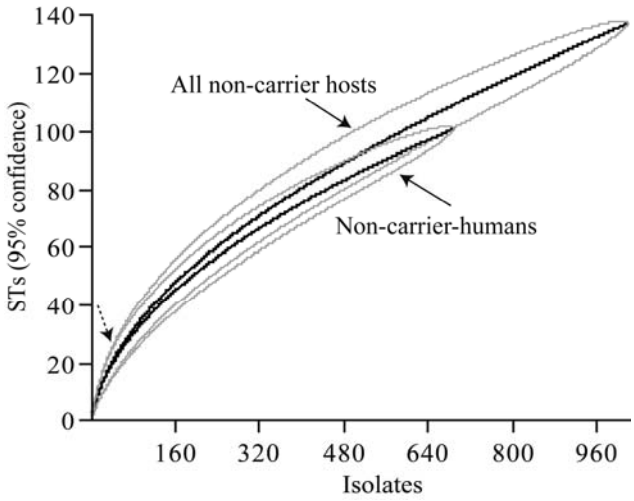

Figure 4.1. Rarefaction curve on isolates from non-carrier humans and ncHA. Number of STs observed among 47 isolates is marked by a dashed arrow.

Table 4.4a. Serovars where all ncHA isolates were the carrier STs

Serovar	Carrier ST (carried isolates)	No. of ncHA isolates
Abony	ST273 (1)	1
Anatum	ST64 (3)	4
Heidelberg	ST15 (2)	15
Panama	ST48 (2)	9
Schwarzengrund	ST96 (1)	2
Thompson	ST26 (1)	10

The STs associated with carriers were the same as for ≥ 50% ncHA isolates within 16 of the remaining 20 serovars (Table 4.4a & b). The exceptions were serovars Derby, Montevideo, Newport and Paratyphi B, where the proportions of ncHA isolates in carrier STs were <40%. Two STs, 33 and 368, were observed among carried isolates of serovar Hadar. ST33 was present in 90% of ncHA isolates but ST368 was not found. Fisher's exact tests were computed for serovars where ≥ 5 ncHA isolates were available to test whether the proportions of carried and ncHA isolates in identical STs were significantly different within each serovar.

The proportions of isolates in the same STs were not significantly different between human carriers and ncHA hosts for any serovar other than Typhimurium (Table 4.4b). All 13 carried isolates of serovar Typhimurium were ST19 whereas only 52% (160/305) of ncHA isolates were in ST19. Other ncHA Typhimurium isolates fell into multiple other STs. Although only 6-31% ncHA isolates shared STs with carried isolates within serovars Derby, Montevideo, Newport and Paratyphi B, the proportions of isolates in shared STs were not significantly different between the two categories for any of these serovars. Only < 5 carried isolates were available for each serovar other than Typhimurium within which 13 carried isolates had been analyzed. And only a single carried isolate was analyzed from 13 of the 22 serovars. These very low numbers of carried isolates may have been responsible for the lack of significance of most comparisons.

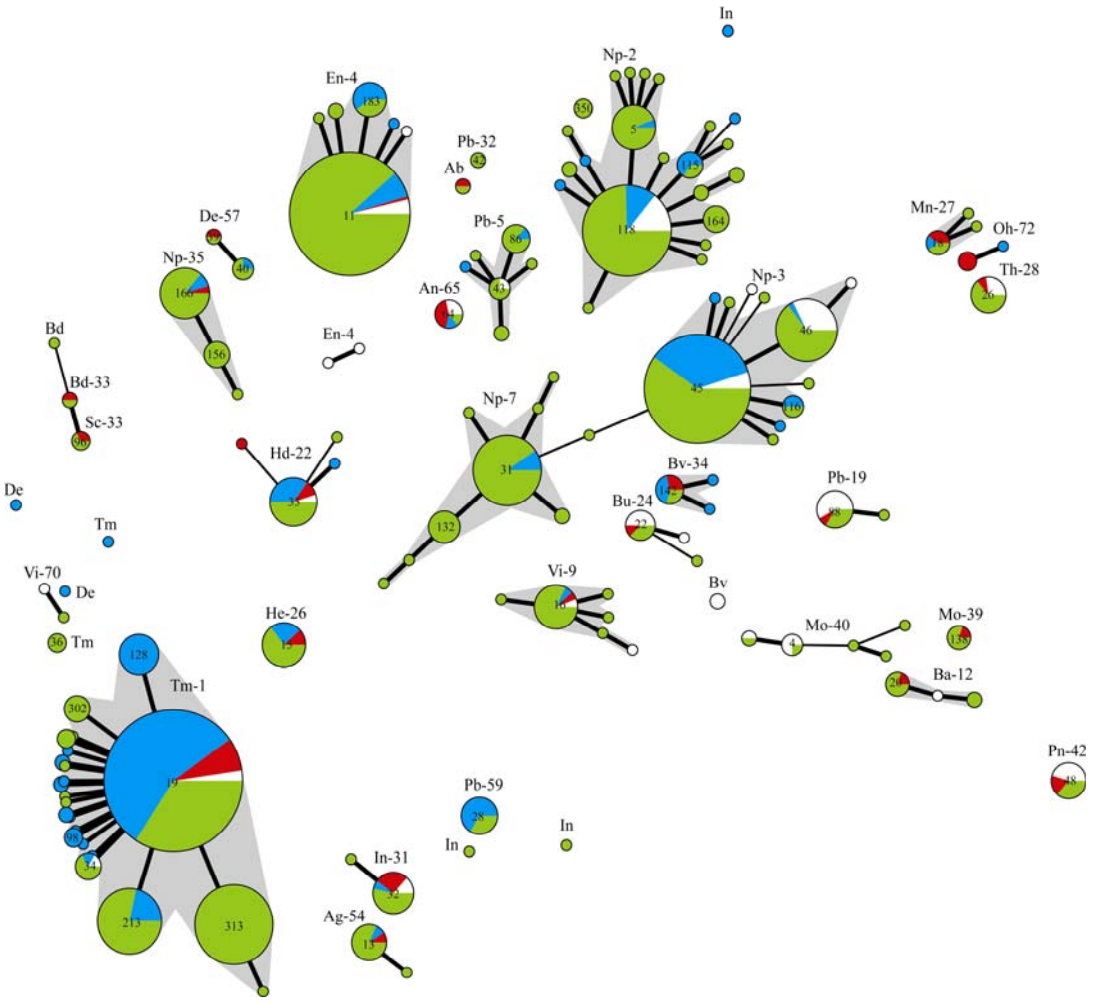

Fig 4.2. An MS_{TREE} of the allelic profiles of isolates. Isolates from carrier humans are shown in red, from non-carrier humans in green, from other mammals in blue and from reptiles in white. Abbreviations used are the mnemonic from Table 4.1 with a suffix that indicates group designation assigned by Prof. Mark Achtman, e.g., Hd-22 is serovar Heidelberg of ST grouping 22.

Table 4.4b. Distribution of ncHA isolates in carrier STs/groups for the remaining serovars

Serovar	Carrier ST (Carried isolates)	Mnemonic-group	No. of ncHA isolates	Proportion[1]	P-value[2]	Proportion[3]
Agona	ST13 (1)	Ag-54	11	0.91 (10)	0.92	1.00 (11)
Bovismorbificans	ST142 (2)	Bv-34	9	0.56 (5)	0.38	0.78 (7)
Braenderup	ST22 (1)	Bu-24	9	0.78 (7)	0.80	1.00 (9)
Brandenburg	ST20 (1)	Ba-12	7	0.57 (4)	0.63	1.00 (7)
Bredeney	ST241 (1)	Bd-33	2	0.50 (1)	n.d.	0.50 (1)
Derby	ST39 (1)	De-57	7	0.14 (1)	0.25	0.71 (5)
Enteritidis	ST11 (1)	En-4	147	0.88 (130)	0.89	1.00 (147)
Hadar	ST33 (2), ST368 (1)	Hd-22	20	0.90 (18)	0.09	1.00 (20)
Infantis	ST32 (4)	In-31	15	0.73 (11)	0.35	0.80 (12)
Manhattan	ST18 (2)	Mn-27	5	0.60 (3)	0.48	1.00 (5)
Montevideo	ST138 (1)	Mo-39	13	0.31 (4)	0.36	0.31 (4)
Newport	ST166 (1)	Np-35	358	0.06 (21)	0.06	0.08 (28)
Ohio	ST329 (3)	Oh-72	1	< 1.00 (0)	n.d.	1.00 (1)
Paratyphi B	ST88 (1)	Pb-19	42	0.26 (11)	0.28	0.29 (12)
Typhimurium	ST19 (13)	Tm-1	305	0.52 (160)	< 0.01	0.99 (301)
Virchow	ST16 (1)	Vi-9	23	0.70 (16)	0.71	0.91 (21)

[1]Proportion of ncHA isolates that were the same ST as carried isolates for respective serovar
[2]Fisher's exact test comparing the proportions of carried and ncHA isolates in the same STs
[3]Proportion of ncHA isolates that shared the groups with carried isolates
n.d., note done where ncHA isolates were < 5 or all isolates shared the group with carried isolates
Note: ST33 and ST368 were observed among carried isolates. The proportions ncHA isolates were independently compared with carried isolates in STs 33 and 368. Both the values were multiplied followed by another multiplication by two.

To further compare the frequencies of occurrence of ncHA isolates in carrier-STs within each serovar, I restricted the analyses to common STs for which at least five ncHA isolates had been identified, thus excluding serovars Abony, Anatum Brandenburg, Derby, Montevideo and Schwarzengrund. Serovars Heidelberg and Panama were also excluded because all ncHA isolates of these serovars were of the same STs as carried isolates. After excluding these serovars, the carrier-ST was the same as an ST containing most of ncHA isolates of the same serovar (Table 4.5). The only exceptions were Newport and Paratyphi B, where carrier-STs were common but not the most common among ncHA isolates.

4.3.4 Prevalence of carrier-STs among ncHA isolates from Europe

Most ncHA isolates had been isolated in Europe except for serovars Newport and Typhimurium where 66% (237/358) and 57% (175/305) of ncHA isolates were from non-European countries. Most ncHA Newport isolates from outside Europe were from the North America (Chapter 3) whereas most non-European Typhimurium were from North America or Africa. The carried serovar Newport isolate was ST166, which is part of lineage

Newport-I that is prevalent among humans in Europe (Chapter 3). Most ST166 isolates were also isolated from European humans. Similarly, ST19, the ST of all carried isolates of serovar Typhimurium was more common among ncHA isolates from Europe than those from outside of Europe ($P < 0.001$). These results suggest that carrier genotypes of serovars Newport and Typhimurium were relatively more prevalent in Europe than other geographic regions.

Table 4.5. Distribution of ncHA isolates in common (≥ 5 isolates) and rare STs (< 5 isolates)

Serovar	Carrier ST	Common STs (isolates)	Rare STs (isolates)	Geo. mean[1]
Agona	13	13 (10)	37 (1)	1.0
Bovismorbificans	142	142 (5)	148 (1), 150 (2), 377 (1)	1.3
Braenderup	22	22 (7)	194 (1), 311 (1)	1.0
Brandenburg	20		20 (4), 65 (1), 249 (2)	2.0
Derby	39		39 (1), 40 (4), 71 (1), 72 (1)	1.4
Enteritidis	11	11 (130), 183 (10)	136 (1), 168(1), 172 (1), 180 (1), 310 (2), 366 (1)	1.1
Hadar	33, 368	33 (18)	12 (1), 123 (1)	1.0
Infantis	32	32 (11)	41(1), 141 (1), 361 (1), 493 (1)	1.0
Montevideo	138		138 (4), 4 (4), 81 (1), 195 (2), 305 (1), 316 (1)	1.8
Newport	166	166 (21), 5 (17), 31 (42), 45(102), 46 (34), 115 (6), 118 (71), 132 (9), 156 (6), 164 (6)	116 (4), 117 (1), 119 (1), 120 (1), 121 (1), 122 (1), 123 (1), 125 (1), 158 (1), 163 (1), 167 (2), 184 (1), 187 (1), 188 (1), 189 (2), 190 (1), 191 (2), 193 (1), 199 (1), 200 (1), 211 (1), 223 (2), 345 (1), 346 (1), 347 (1), 348 (1), 349 (1), 350 (3), 351 (1), 352 (1), 353 (1), 354 (1), 355 (1), 360 (1)	1.2
Paratyphi B	88	88 (11), 28 (12), 86 (7)	42 (2), 43 (4), 110 (2), 127 (1), 149 (1), 307 (1), 570 (1)	1.5
Typhimurium	19	19 (160), 34 (6), 128 (14), 213 (37), 302 (6), 313 (55)	35 (1), 36 (3), 98 (3), 99 (1), 137 (1), 153 (1), 204 (1), 205 (1), 207(1), 209 (1), 323 (2), 328 (1), 332 (2), 376 (2), 394 (1), 429 (1), 568 (3), 569 (1)	1.3
Virchow	16	16 (16)	38 (1), 181 (1), 197 (1), 303 (1), 326 (1), 333 (1), 359 (1)	1.0

[1]Geometric mean of the isolates in rare STs

4.3.5 Carrier-STs in groups with ncHA isolates

The MLST scheme used in this study has been widely accepted by the *Salmonella* community, leading to MLST data for > 2350 *Salmonella* isolates on the website (http://web.mpiib-berlin.mpg.de/mlst/dbs/Senterica) (June 2008). Ninety groups were identified among these isolates by Prof. Mark Achtman on the basis of phylogenetic analyses of nucleotide sequences and an MS_{TREE} of allelic profiles (unpublished data). Each

phylogenetic group of ≥ 2 STs and all distinct STs with ≥ 4 isolates were assigned a unique group designation and represents a distinct cluster of genetically closely related isolates.

For most serovars, multiple isolates possessed STs that differed from the carrier-STs (Table 4.4b, Table 4.5). However, it was not clear whether isolates in those STs were genetically distant from carried isolates. Therefore, I used the groups assigned by Prof. Achtman to calculate the proportion of ncHA isolates that was genetically related to the carried isolates within each serovar.

For six serovars, all ncHA isolates were of the same ST as the carrier-STs (Table 4.4a). Furthermore, all ncHA isolates of serovars Agona, Braenderup, Brandenburg, Enteritidis, Hadar and Manhattan were in the same groups as carried isolates of those serovars (Table 4.4b). A high proportion of other isolates were also in the same groups as carrier-STs and only 1-29% isolates were genetically distinct. However, for serovars Montevideo, Newport and Paratyphi B, the groups containing carrier STs were rarely found among ncHA isolates, indicating that most isolates in these serovars are genetically distinct from carried isolates (Table 4.4b). In serovar Typhimurium, only 52% of ncHA isolates were in ST19 but 99% of ncHA isolates were in the same group (Group 1) as carrier isolates. These results indicate that most of ncHA isolates were genetically closely related to the carried isolates in most serovars.

4.4 Discussion

4.4.1 Diversity between isolates from asymptomatic human carriers and non-carrier humans and animals

The genetic diversity was higher among *Neisseria meningitidis* isolates from asymptomatic carriers than clinical isolates, both by MLEE and MLST analyses (Caugant *et al.*, 1988; Yazdankhah *et al.*, 2004). In contrast, *Salmonella* isolates from asymptomatic human carriers showed similar levels of nucleotide diversity and genotype richness as non-carrier isolates of the same serovars from humans and animals (Fig 4.1, Table 4.2). These observations possibly reflect a difference in the biology of carrier and non-carrier isolates between the two species. Multiple unique alleles were found that distinguished carried from invasive *N. meningitidis* (Jolley *et al.*, 2000) but all the alleles in carried salmonellae were common among ncHA isolates. Furthermore, several groups of STs were preferentially associated with disease in *N. meningitidis*, indicating a differential association between clonal lineages of meningococci and disease (Yazdankhah *et al.*, 2004). In contrast, most

Salmonella STs found among isolates from carriers were common to ncHA isolates within each serovar and with similar frequencies (Table 4.4a & b). The proportions of ncHA isolates in the groups with carried isolates were even higher than those in carrier-STs for most serovars (Table 4.4b). These observations indicate that *Salmonella* isolates from human carriers probably belong to the same population as isolates from non-carrier humans and animals. However, only few isolates from human carriers were tested except for serovar Typhimurium, which may be responsible for the lack of significance of most comparisons.

The results of my study are in general agreement with the previous study that found a comparable overall genotypic diversity among Typhimurium DT104 isolates from asymptomatic and sick pigs by pulsed-field gel electrophoresis (Perron *et al.*, 2007). However, isolates from asymptomatic pigs were more diverse within individual herds, leading to the suggestion that selection bias was followed by selective sweep of adaptive genotypes that resulted in lower genotypic diversity in disease associated populations within herds (Perron *et al.*, 2007).

4.4.2 Prevalence of carrier genotypes among non-carrier isolates

The carrier-STs were generally the most common among ncHA isolates for most serovars (Table 4.4a & 4.4b and Table 4.5). >50% ncHA isolates were the same STs as carried isolates of most serovars, except for serovars Derby, Montevideo, Newport and Paratyphi B where only 6-31% of ncHA isolates shared STs with carried isolates. The proportions of carried and ncHA isolates in the same ST were similar for all serovars except for Typhimurium (Table 4.4b), where only 52% ncHA isolates were the same ST as carried isolates. These results indicate the existence of common genotypes for most serovars that can asymptomatically inhabit the hepatobiliary system in human carriers as well as can cause disease in other individuals.

All carried isolates and a high proportion of ncHA isolates were from Europe, except for serovars Newport and Typhimurium where ncHA isolates from outside Europe were most frequent. The carrier-ST of serovar Newport was more prevalent within Europe because only one of the 21 ncHA isolates in this ST was from outside Europe (Chapter 3). Similarly, the carrier genotype was more frequent among ncHA isolates of serovar Typhimurium from Europe than other geographic regions ($P < 0.001$). Possibly these carrier genotypes are more prevalent in Europe than elsewhere. However, more ncHA

isolates from other geographic regions would need to be analyzed for the remaining serovars to confirm this observation.

The carrier-ST was the common ST for most serovars (Table 4.5). For 12 serovars, all ncHA isolates either shared the carrier-STs or were in the same groups as the carried isolates (Fig 4.2, Table 4.4a & 4.4b). The proportion of ncHA isolates in the same groups as carried isolates was higher (71-99%) than in other groups or STs for the remaining serovars, except for Montevideo, Newport and Paratyphi B where most isolates were genetically distinct from carrier-STs. These results suggest that *Salmonella* isolates from chronic human carriers are genetically identical to isolates from non-carrier humans and animals within most serovar. However, all ncHA isolates of a serovar are not necessarily the same as carried isolates.

5. Clonal diversity and host association in subspecies *enterica*

5.1 Introduction

Serovars of *S. enterica* vary greatly in their host range (Kingsley and Baumler, 2000). Most serovars are generalists and have a broad host range, e.g. Enteritidis and Typhimurium (Uzzau *et al.*, 2000). In contrast, some serovars are host restricted and only infect specific hosts, e.g., Typhi that can only infect humans and higher primates (Kingsley and Baumler, 2000; Uzzau *et al.*, 2000). Still other serovars are adapted to particular hosts but are occasionally isolated from other sources, e.g. Choleraesuis and Dublin that are adapted to pigs and cattle, respectively (Kingsley and Baumler, 2000; Uzzau *et al.*, 2000). These are called host adapted.

Host restricted and host adapted variants have also been identified within some generalist serovars. Phage typing has identified some host associated subtypes within serovars Typhimurium and Enteritidis (Hickman-Brenner *et al.*, 1991; Rabsch *et al.*, 2002; van Duijkeren *et al.*, 2002). DT2 and DT99 are two Typhimurium phage types that are restricted to pigeons (Rabsch *et al.*, 2002). Enteritidis phage types, PT4 and PT11, are associated with poultry and hedgehogs, respectively (Coyle *et al.*, 1988; Nauerby *et al.*, 2000; Riley and Chomel, 2005). Host associated lineages have also been identified for some *S. enterica* serovars by other approaches that are based on selectively neutral loci and which have been extensively used in population genetic studies. Based on the relative frequencies of isolates from various hosts, MLEE analysis has identified bird and mammal associated lineages of Derby as well as human and animal associated lineages of Newport (Beltran *et al.*, 1988). The occurrence of host associated subtypes has also been demonstrated within subspecies *enterica* by MLST (Alcaine *et al.*, 2006). A collection of 156 bovine and 179 human isolates belonging to 52 serovars was analyzed by a three gene MLST scheme and STs that were unique to bovine or human isolates were observed in addition to STs present in diverse hosts (Alcaine *et al.*, 2006). Human and bovine associated lineages were also identified within Newport by this scheme (Alcaine *et al.*, 2006; Sukhnanand *et al.*, 2005). Finally, host associated subtypes were identified within serovar Enteritidis when a diverse collection of 34 isolates was analyzed by MLVA (Cho *et al.*, 2007).

Although reptiles are the primary reservoir for *Salmonella*, and reptile-associated salmonellosis is a global problem (Ebani *et al.*, 2005; Geue and Loschner, 2002; Nakadai *et*

al., 2005; Centers for Disease Control and Prevention, 2007c; Woodward *et al.*, 1997), most studies have been focused on isolates from human and animals. In this study, reptile isolates of subspecies *enterica* were analyzed by an MLST scheme that is based on seven housekeeping genes (Torpdahl *et al.*, 2005). In order to compare their host associations within serovars, I only tested reptile isolates from serovars where MLST data were available for isolates from humans or non-human warm blooded animals (NhWBA). 137 reptile isolates from 23 serovars matched these criteria and were tested by MLST.

5.2 Methods

5.2.1 Bacterial isolates

109 isolates that had been isolated from reptiles in different parts of Germany were obtained from Dr. Reiner Helmuth, National Salmonella Reference Laboratory, Federal Institute for Risk Assessment (BFR), Berlin, Germany. 28 reptile isolates of serovar Newport from Chapter 3 were also included. In total, these 137 isolates belonged to 23 serovars of subspecies *enterica.*

MLST data were available for 748 isolates for the same 23 serovars from humans and 269 isolates from NhWBA (birds and other mammals) on the *Salmonella* MLST website. This dataset includes 268 Newport isolates from humans and 69 isolates from NhWBA from Chapter 3. MLST data for host adapted serovar Choleraesuis (74 isolates) and host restricted serovars Paratyphi A (41 isolates), Paratyphi C (48 isolates) and Typhi (27 isolates) were also downloaded for comparison with any host adapted variants of these generalist serovars. Details are summarized in Table 5.1.

The cultivation and storage of isolates, isolation of DNA, PCR amplification, sequencing and MLST analyses were carried out as previously described (Chapter 3). An MS_{TREE} was generated in Bionumerics 4.5 based on the allelic profiles of isolates.

5.2.2 Characteristics of housekeeping genes

The average pairwise nucleotide diversity (π) with Jukes-Cantor correction was calculated on concatenated sequence alignments of isolates from reptiles, humans and NhWBA using MEGA 4.0 (Tamura *et al.*, 2007). The mean of non-synonymous substitutions per non-

synonymous site (Ka) and synonymous substitutions per synonymous site (Ks) was calculated for alignments of gene fragments in each host category using DnaSP 4.0 (Rozas *et al.*, 2003). Unique sequences from each of the seven gene fragments present in any host category were tested for positive selection using the codon substitution models implemented in PAML 4.0 (Yang, 2007). Likelihood ratio tests were computed between the codon substitution models to find the model with the best fit (Yang *et al.*, 2000). Only sites that were inferred to be under positive selection with a posterior probability of > 95% were considered.

Table 5.1. Details of isolates from each host category

Serovars; Mnemonic	Number of isolates		
	Reptiles	Humans	NhWBA[1]
Anatum; An	2	4	1
Bovismorbificans; Bv	2	4	5
Braenderup; Bu	5	5	
Brandenburg; Ba	1	7	
Choleraesuis[2]; Cs		37	37
Decatur; Dc	1		2
Dublin; Du	2	11	9
Enteritidis; En	8	124	16
Hadar; Hd	1	13	8
Infantis; In	2	15	2
Javiana[3]; Jv	4	6	
Miami; Mi	2	5	
Montevideo; Mo	4	10	
Muenchen; Mu	15	7	
Newport; Np	28	268	62
Oranienburg; Or	30	5	
Panama; Pn	5	6	
Paratyphi A; Pa		41	
Paratyphi B[4]; Pb	8	25	10
Paratyphi C[5]; Pc		48	
Saintpaul; Sp	2	12	14
Senftenberg; Sf	3	10	2
Stanley; St	1	4	
Thompson; Th	3	8	
Typhi; Ty		27	
Typhimurium; Tm	5	179	137
Virchow; Vi	3	20	1
Total	137	901	306

[1]Isolates from non-human warm blooded animals (NhWBA)
[2]Including Choleraesuis var. Kunzendorf
[3]Including two z28 negative (H1 antigen) isolates from reptiles
[4]Including Paratyphi B var. Java from humans and NhWBA
[5]Including Paratyphi C var. Hirschfeld, Orient and Orientalis

5.2.3 Rarefaction curves

Rarefaction is one of the best methods to compare the number of taxa between samples of different sizes because it allows estimating the number of taxa to be expected for a smaller sample from the data for larger populations. Individual based rarefaction curves were computed on isolates from reptiles, humans and NhWBA using the paleontological statistics software package PAST, version 1.73 (Hammer *et al.*, 2001).

Isolates of serovars Enteritidis, Newport and Typhimurium from humans and NhWBA and serovar Oranienburg from reptiles were overrepresented in the dataset. Therefore, I re-computed rarefaction curves on isolates from each host category after excluding these serovars.

5.2.4 Simpson's diversity index (1-D)

Simpson's diversity index (SID) is often used to quantify the species diversity within a habitat. SID (1-D) indicates the probability that two randomly selected individuals from a sample will be from different species, where $D = \Sigma (n / N)^2$. In this formula, n is the number of individuals of a particular species and N is the total number of individuals in the sample. SID was calculated using PAST version 1.73 (Hammer *et al.*, 2001) for isolates from reptiles, humans and NhWBA. SID was also calculated for these host categories after excluding serovars Enteritidis, Newport, Oranienburg and Typhimurium.

5.2.5 The UniFrac test of significance and the P-test

Given a phylogenetic tree and environmental information for each sequence, the UniFrac test of significance computes differences between communities as a fraction of the branch lengths that are unique to each environment i.e. that leads to the descendents of either of environment but not from both (Lozupone and Knight, 2005). The program calculates a probability, P, that corresponds to the fraction of permutations where more branch lengths are unique in the original tree than in a tree where environment assignments were randomized (Lozupone *et al.*, 2006).

The program UniFrac also computes a P-test that compares communities by counting the numbers of changes responsible for the distribution of sequences from different environments in a phylogenetic tree (Martin, 2002). The probability is calculated as the fraction of permutated trees with randomized environments that needed fewer changes to describe the distribution than the original tree.

PAUP package 4.0 (Swofford, 1998) was used to generate a Neighbor-joining (NJ) tree on a concatenated sequence alignment for the seven housekeeping genes using only unique sequences from each host category. The UniFrac tests of significance were computed with 1,000 permutations without weight abundance on all environments together and on each pair of environments. The P-tests were also computed with 1,000 permutations on all environments together and on each pair of environments.

5.2.6 Other statistical tests

Fisher's exact tests and χ^2 tests were computed using the statistical package STATISTICA (StatSoft, Inc., Tulsa, U.S.A.).

5.3 Results

5.3.1 Characteristics of housekeeping genes

An adaptation to a particular ecological niche may involve evolutionary pressure which might possibly be reflected by variable nucleotide diversity between the individuals of a species from different environments. Therefore, I compared nucleotide diversity (π) of concatenated sequences between isolates from each host category. Serovars Choleraesuis, Paratyphi A, Paratyphi C and Typhi were not included because isolates of these serovars were not available from reptiles. The π values were comparable between reptile and human isolates but π was lower among isolates from NhWBA (Table 5.2). Hence, nucleotide diversity did not indicate any difference in selection pressure between isolates from reptiles and humans. A difference between the π values of isolates from reptiles and NhWBA might indicate a difference in the selection pressure or a selective advantage of particular sequence types in NhWBA hosts.

To test whether housekeeping genes in any host category have experienced positive selection, ω, the ratio of nonsynonymous substitutions per non-synonymous site (Ka) to synonymous substitution per synonymous site (Ks) was calculated for sequence alignments of gene fragments (Table 5.3). The ω (Ka/Ks) values were < 1.0, as expected for purifying selection (Perez-Losada *et al.*, 2006), indicating that none of the gene fragments was under positive selection in any of the three host categories. However, ω was 4-10fold higher for the *hemD* gene fragment than for other gene fragments, indicating that non-synonymous substitutions were fixed at higher frequencies in *hemD* than other fragments in all host categories.

Table 5.2. Diversity within each host category

Host category	Nucleotide diversity	Rarefied STs[1]	SID (1-D)[2]	Rarefied STs[3]	SID (1-D)[4]
Reptile	0.012 ± 0.001	47	0.96	28.3 ± 1.4	0.94
Human	0.010 ± 0.001	45.2 ± 3.8	0.94	30.6 ± 2.5	0.96
NhWBA	0.007 ± 0.001	33.7 ± 2.7	0.84	18	0.87

Note: isolates of serovars Choleraesuis, Paratyphi A, Paratyphi C and Typhi were not included because these serovars are not isolated from reptiles.
[1]Number of STs identified by rarefaction analysis for 137 isolates against 47 STs of reptile isolates
[2]SID on all isolates in each host category
[3]Number of STs identified by rarefaction analysis for 54 isolates against 18 STs of NhWBA isolates after excluding isolates of serovars Enteritidis, Newport, Oranienburg and Typhimurium
[4]SID in each host category after excluding isolates of serovars Enteritidis, Newport, Oranienburg and Typhimurium.

Table 5.3. Average pairwise distance at non-synonymous (Ka) and synonymous (Ks) sites among gene fragments in different host categories

Gene fragments	Reptile			Human			NhWBA		
	Ka	Ks	ω	Ka	Ks	ω	Ka	Ks	ω
aroC	0.00016	0.02712	0.0059	0.00014	0.02610	0.0054	<0.00002	0.01509	<0.0013
dnaN	0.00081	0.04208	0.0192	0.00129	0.02539	0.0508	0.00122	0.01705	0.0716
hemD	0.00434	0.02259	0.1921	0.00488	0.02195	0.2223	0.00481	0.01610	0.2987
hisD	0.00101	0.05847	0.0173	0.00223	0.04854	0.0459	0.00197	0.02979	0.0661
purE	0.00034	0.05120	0.0066	0.00101	0.03323	0.0304	0.00022	0.02319	0.0095
sucA	0.00008	0.04397	0.0018	0.00002	0.03364	0.0006	0.00028	0.02519	0.0111
thrA	0.00004	0.06469	0.0006	0.00004	0.06722	0.0006	0.00004	0.05517	0.0007

Note: Isolates of serovars Choleraesuis, Paratyphi A, Paratyphi C and Typhi were not included.

To confirm that the seven gene fragments were not under positive selection, alignments of all unique fragments for each gene from all three host categories were also tested using various codon substitution models implemented in PAML 4.0. In agreement to the ω values, none of the gene fragments is under overall positive selection. At the single codon level, only one codon in *aroC* (position 110), four in *hemD* (positions 33, 34, 114 and 115) and one in *hisD* (position 6) are under diversifying selection at the > 95% level according to the M2 (positive selection), M3 (discrete site) or M8 (β and ω) models. None of the models differ significantly from the others according to the likelihood ratio test for any gene fragment. The only exception was model M1 (nearly neutral) for the *purE* gene fragment where the likelihood was marginally lower than M2 and M3.

5.3.2 Diversity between different host categories

A total of 47 STs were observed among the 137 reptile isolates. Since more isolates had been tested from humans and NhWBA, the numbers of STs that would be expected for 137 isolates from humans and NhWBA were estimated using individual rarefaction. Again,

isolates of serovars Choleraesuis, Paratyphi A, Paratyphi C and Typhi were not included. The expected value for the number of STs among human isolates was comparable to the number of STs observed among reptile isolates (Fig. 5.1, Table 5.2). The genotypic diversity among isolates from NhWBA was much lower because only 33.7 ± 2.7 STs would be expected for 137 isolates.

The Simpson's Index of Diversity (SID) was also comparable between reptile and human isolates. Similar to the rarefaction analysis, SID was lower for isolates from NhWBA (Table 5.2). Isolates of serovars Enteritidis, Newport and Typhimurium from humans and NhWBA were overrepresented in the dataset because other projects had focussed on these serovars. Isolates of serovar Oranienburg from reptiles (n = 30) were also 6fold higher in numbers than from other hosts (n = 5). I therefore calculated rarefaction and SID values again, but after excluding isolates of these four serovars from all host categories. Now the smallest category was NhWBA (54 isolates, 18 STs), and therefore rarefaction values were estimated for 54 isolates each from reptiles and humans. The results were consistent with the previous observations. The rarefied STs and SID values were comparable between reptile and human isolates but higher than for isolates from NhWBA (Table 5.2). A lower diversity among NhWBA isolates possibly indicates selective advantages for a limited number of genotypes in the host environment.

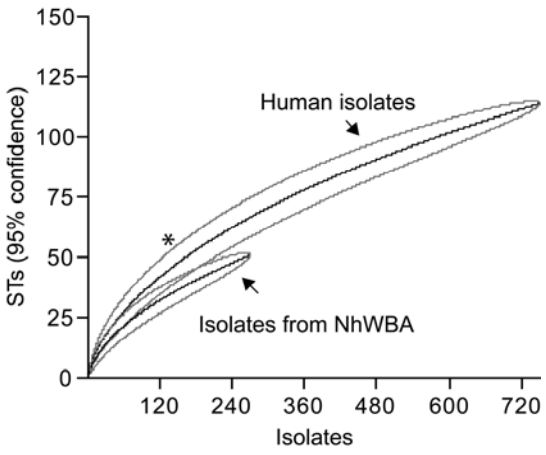

Fig 5.1. Rarefaction curves on isolates from humans and NhWBA. STs for 137 isolates are marked with a star sign.

5.3.3 Host-unique sequence types (STs) among generalist salmonellae

An MS_{TREE} generated from the allelic profile of isolates (Fig. 5.2) defined STs that were common between the host categories (multi-coloured pie-charts) as well as STs that were specific to a particular host type (uni-coloured) within generalist serovars. Of the 47 STs observed among reptile isolates, 12 were present in all three host categories (Fig. 5.3A). An

additional 12 STs were shared by reptile and human isolates. Reptile and NhWBA isolates did not share any unique STs. Approximately 49%, 68% and 51% of the STs were specific for reptile, human and NhWBA isolates, respectively. Isolates of host restricted and host adapted serovars (Choleraesuis, Paratyphi A, Paratyphi C and Typhi) were excluded from these comparisons.

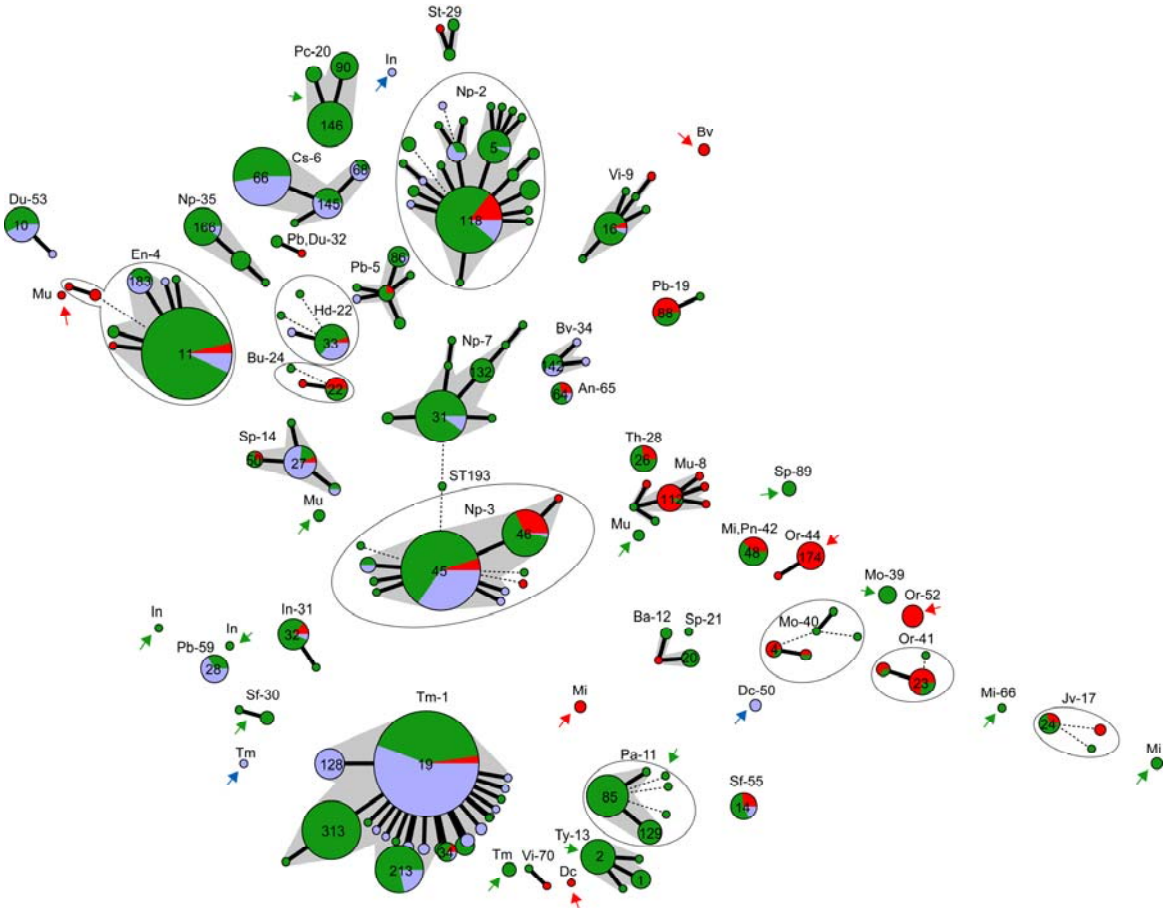

Fig. 5.2. An MS_{TREE} of allelic profiles of isolates. Reptile isolates are shown in red, human isolates in green and NhWBA in blue. Abbreviations used as mentioned in Table 5.1, followed by group designation assigned by Prof. Mark Achtman. Host specific groups or STs are marked with arrows. ST193, a Newport ST, is a DLV to ST45 of group 3 and ST31 of group 7.

The proportion of reptile isolates in host-specific STs was slightly but insignificantly ($p = 0.08$) higher than for human isolates and significantly higher than for NhWBA isolates ($p < 0.01$). The proportion of NhWBA isolates in host-specific STs was also significantly lower than human isolates in human-specific STs ($p < 0.01$). These results indicate that isolates from reptiles are more often host-specific than isolates from humans or NhWBA. However, a high proportion of Oranienburg isolates from reptiles might partially be responsible for such a phenomenon. 67% of reptile isolates of serovar Oranienburg were present in host-

specific STs. A significantly lower proportion of NhWBA isolates in host-specific STs possibly indicates that NhWBA-specific isolates are rare.

Table 5.4. Distribution of STs observed for the isolates in different host-categories

Host	No. of observed STs (Total isolates)	Common STs (% isolates)	Host specific STs (% isolates)	Host specific groups (% isolates)
Reptile	47 (137)	24 (67.9)	23 (32.1)	6 (19.0)
Human	114 (748)	36 (75.1)	78 (24.9)	10 (3.0)
NhWBA	49 (269)	24 (83.6)	25 (16.4)	3 (1.5)

Note Isolates of serovans Choleraesuis, Paratyphi A, Paratyphi C and Typhi were not included.

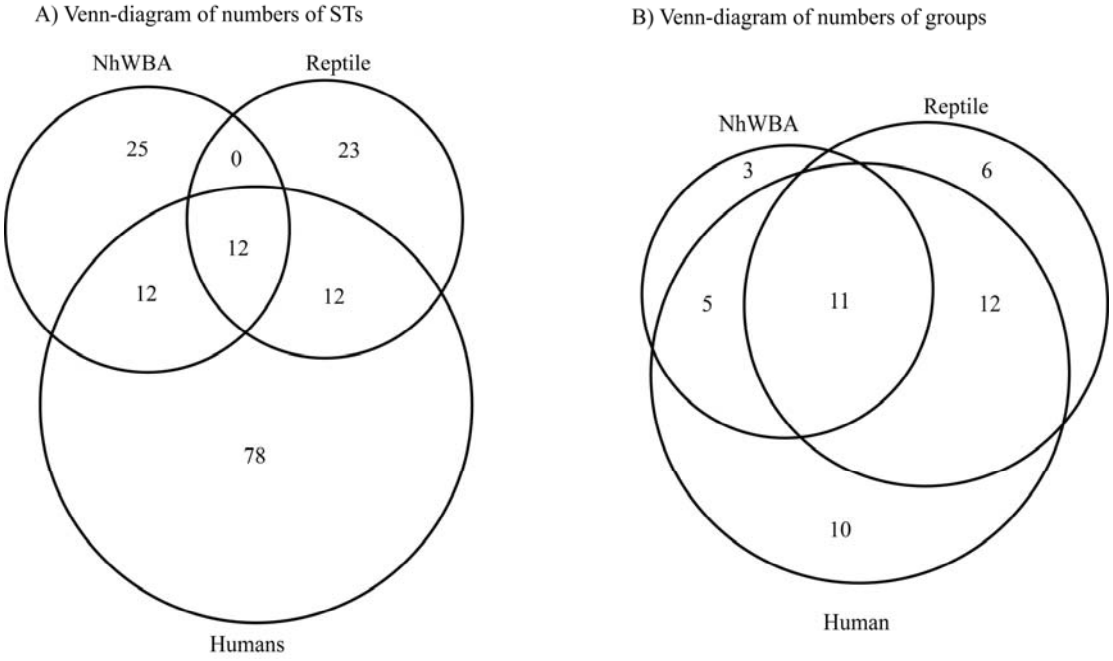

Fig 5.3. Proportional Venn-diagrams A). the number of STs shared between the three host categories, and B). the numbers of groups shared between the host categories.

5.3.4 Host restricted groups among generalist salmonellae

My analyses of salmonellae from human carriers revealed that most STs found among carried isolates were also found among isolates from human non-carriers and from animals (Chapter 4). Only one ST in serovar Hadar was specific to carriage. Other STs were only observed among isolates from non-carrier humans and from animals but most were genetically closely related to carrier isolates. In contrast, approximately half (23/47) of the STs among reptile isolates were specific to reptiles. To test whether STs unique to reptile

hosts represent genetically distinct populations, the groups defined by Prof. Mark Achtman were used (Fig. 5.1, Fig. 5.3B, Table 5.5).

Table 5.5. Distribution of groups according to serovar and host specificity

Serovar	Shared groups [Hosts (ST designation)]	Host specific groups [Hosts (ST designation)] or singleton
Anatum	**65** [R-H-N (64)]	
Bovismorbificans	**34** [H-N (142), N (148, 377)]	[R (150)]
Braenderup	**24** [R-H (22), R (194), H (311)]	
Brandenburg	**12** [R (65), H (20, 249)]	
Choleraesuis	**6** [H-N (66, 68, 145), H (133)]	
Decatur		**50** [N (70)]; [R (186)]
Dublin	**32** [R (74)[1]]; **53** [H-N (10), N (73)]; **4** [R (180)[2]]	
Enteritidis	**4** [R-H-N (11), H-N (183), R (168), H (310, 366), N (136), R (172, 180)[2]]	
Hadar	**22** [R-H-N (33), N (12), H (368, 473)]	
Infantis	**31** [R-H-N (32), H (41)]	[H (361)]; [H(493)]; [N (141)]
Javiana	**17** [R-H (24), R (175), H (143)]	
Miami	**42** [H (48)[3]]	**66** [H (140)[4]]; [R (171)]; [H (80)]
Montevideo	**40** [R-H (4), R-H (195), H (81, 305, 316)]	**39** [H (138)]
Muenchen	**8** [R-H (112), R (170, 173, 176, 177), H (82, 111)]	[R (178)]; [H (83)]; [H(84)]
Newport	**2** [R-H-N (118), H-N (5, 115), H (117, 119, 163, 164, 167, 187, 189, 190, 199, 223, 345, 347, 350, 351, 352, 354), N (120, 122, 123)], **3** [R-H-N (45, 46,), H-N (116), H (353, 355, 158, 201), N (121, 125), R (184, 211)], **7** [H-N (31), H (132, 188, 191, 200, 346, 348, 349)], **35** [H-N (166), H (156, 360)]; ST193[5]	
Oranienburg	**41** [R-H (23, 47), H (91)]	**44** [R (169, 174)]; **52** [R (179)]
Panama	**42** [R-H (48)[3]]	
Paratyphi A		**11** [H (85, 129, 130, 479, 494, 495)]
Paratyphi B (including var. java)	**5** [R-H (43), H-N (86), H (110, 307, 570), N (149)]; **19** [R-H (88), H (127)]; **32** [H (42)[1]]; **59** [H-N (28)]	
Paratyphi C		**20** [H (90, 114, 146)]
Saintpaul	**14** [R-H-N (27), R-H (50), H-N (49), H (344)]; **21** [H (343)[6]]	**89** [H (95)]
Senftenberg	**55** [R-H-N (14)]	**30** [H (185, 217)]
Stanley	**29** [R (182), H (29, 51)]	
Thompson	**28** [R-H (26)]	
Typhi		**13** [H (1, 2, 3, 8)]
Typhimurium	**1** [R-H-N (19, 34), H-N (213), H (35, 137, 302, 313, 328, 394, 568, 569), N (98, 99, 128, 153, 204, 205, 209, 323, 332, 376, 429)]	[H (36)]; [N (207)]
Virchow	**9** [R-H-N (16), H (38, 326, 303, 359)]; **70** [H (333), R (197)]	

R, reptile host; H, human host; N, non-human warm blooded animal host
Note: 1.Host categories separated by hyphens before an ST in parentheses indicate that that ST is common to them.
2. The group designations are in bold face.
3. The groups and/or the distinct STs in each category are separated by semicolons.

Table 5.5 continued…

[1]Group 32 contained two STs, ST74 which was unique to a reptile isolate of serovar Dublin and ST42 which was specific to two human isolates of Paratyphi B.
[2]ST180 is a part of group 4 and most group 4 isolates are Entertidis. ST180 was common to reptile isolates of serovar Dublin and Enteritidis.
[3]Group 42, a single ST group, is common between serovars Miami and Panama.
[4]Group 66 contained STs 140 and 410. The former is specific to a human isolate of serovar Miami whereas the latter was observed for an isolate of serovar Eastbourne of unknown origin.
[5]ST193 is a DLV between ST45 of group 3 and ST31 of group 7.
[6]ST343 was assigned to group 21. Group 21 has two other STs 126 and 412. Serovar of the sole isolate in ST126 was not determined whereas ST412 contained an isolate of serovar Reading.

All isolates of serovars Anatum, Braenderup, Brandenburg, Enteritidis, Hadar, Javiana, Panama, Stanley and Thompson belonged to a single serovar-specific group (Fig. 5.2, Table 5.5). Two or more distinct groups were identified within serovars Dublin, Newport, Paratyphi B and Virchow but none of them was specific to particular host. Host-specific groups or singleton STs were observed only in serovars Bovismorbificans, Decatur, Infantis, Miami, Montevideo, Muenchen, Oranienburg, Saintpaul, Senftenberg and Typhimurium (Table 5.5).

Two group and four singletons that were specific to reptile isolates (Table 5.6) are marked with red arrows in Fig. 5.2. Four groups plus six singletons specific to isolates from humans (green arrows in Fig. 5.2) and one group plus two singletons were specific to NhWBA (blue arrows in Fig. 5.2). Host adapted and restricted serovars Choleraesuis, Paratyphi A, Paratyphi C and Typhi were also excluded from these comparisons. The proportion of reptile isolates in host-specific groups and singletons was significantly higher than for isolates from humans ($p < 0.01$) or NhWBA ($p < 0.01$). However, host-specific groups and singletons for human and NhWBA isolates contained similar proportions of isolates ($p = 0.14$). Therefore, in agreement with the previous observation, host specific salmonellae are relatively more common among reptiles than humans or NhWBA.

To test whether reptiles represent a distinct host community from humans and NhWBA, the UniFrac significance test and the P-test were computed. The host categories did not differentiate the isolates as indicated by the UniFrac significance test ($p = 0.72$) or the P-test ($p = 0.22$) when all host categories were compared together. Pairwise comparisons also did not resolve any significant differences between the host categories by either method (Table 5.7). These results indicate that generalist salmonellae from reptiles are not divergent from those isolated from humans or NhWBA and belong to the same population.

Table 5.6. Host specific groups and STs

Host	Serovar	Group [ST]	Number of isolates
Reptile	Bovismorbificans	[150]	2
	Decatur	[186]	1
	Miami	[171]	2
	Muenchen	[178]	1
	Oranienburg	**44** [169, 174]	13
	Oranienburg	**52** [179]	7
Human	Infantis	[361]	1
	Infantis	[493]	1
	Miami	**66** [140]	1
	Miami	[80]	2
	Montevideo	**39** [138]	5
	Muenchen	[83]	2
	Muenchen	[84]	2
	Paratyphi A	**11** [85, 129, 130, 479, 494, 495]	41
	Paratyphi C	**20** [90, 114, 140]	48
	Saintpaul	**89** [95]	3
	Senftenberg	**30** [185, 217]	4
	Typhi	**13** [1, 2, 3, 8]	27
	Typhimurium	[36]	3
NhWBA	Decatur	**50** [70]	2
	Infantis	[141]	1
	Typhimurium	[207]	1

Table 5.7. Probability matrix of pairwise comparison of host categories

	Reptile	Human	NhWBA
Reptile		1.0000	0.6600
Human	1.0000		1.0000
NhWBA	0.0660	0.6600	

Note: 1. Probabilities of the UniFrac test of significance are shown in the upper half and the P-test probabilities in the lower half of the matrix. 2. Isolates of serovars Choleraesuis, Paratyphi A, Paratyphi C and Typhi were not included in the analysis.

5.3.5 Host-adapted serovars/variants and host specific genotypes

All human and cattle isolates of cattle adapted serovar Dublin clustered together in group 53 (Fig. 5.2, Table 5.5, Table 5.8). Similar results were obtained for the swine adapted serovar Choleraesuis, where all human and swine isolates were in group 6. Thus, human cases of Dublin and Choleraesuis infections probably originate from cattle and swine, respectively.

I also analyzed poultry-associated phage type PT4 and hedgehog-associated PT11 of serovar Enteritidis and pigeon-restricted phage types DT2 and DT99 of serovar Typhimurium. All isolates of each phage type were confined to a single ST within groups

that were common to isolates from reptiles, humans as well as other animals (Table 5.8). One Typhimurium DT99 isolate was the exception that was in ST128 whereas the remaining 10 DT99 isolates were in ST19. None of these STs was host specific. These patterns are different from those of host restricted serovars Paratyphi A, Paratyphi C and Typhi in which all isolates of these serovars were assigned to a single host- and serovar-specific group of STs (Fig. 5.2, Table 5.6). Thus, Typhimurium DT2 and DT99 are apparently not totally host-restricted. Enteritidis PT4 and PT11 are not host-restricted. All 13 Enteritidis PT4 isolates that were analyzed in this study were from humans whereas PT11 isolates were available both from hedgehogs (n = 5) and humans (n = 4). All human as well as hedgehog isolates of Enteritidis PT11 were the same ST (ST183). Therefore, similar to the host adapted serovars, Enteritidis PT11 infections in humans are probably derived from hedgehogs.

Table 5.8. Host-specificity among host-associated serovars/variants

Serovar (Phage type)	Primary host (isolates)	Group [STs]	Isolates from other hosts in the same group		
			Reptile	Human	Other
Choleraesuis	Swine (37)	**6** [66, 68, 145]	0	37	0
Dublin	Cattle (6)	**53** [10, 73]	0	11	3
Enteritidis (PT4, 4b)	Poultry (0)	**4** [11]	8	124	16
Enteritidis (PT11)	Hedgehogs (5)	**4** [183]	8	124	7
Typhimurium (DT2)	Pigeon (12)	**1** [128]	4	176	114
Typhimurium (DT99)	Pigeon (11)	**1** [19, 128]	4	176	114

Note: A total of 11 Enteritidis PT4 and 2 PT4b isolates were analyzed but all were isolated from humans.

5.4 Discussion

5.4.1 Genetic and genotypic diversity

Asymptomatic isolates of *Neisseria meningitidis* were more diverse than clinical isolates (Caugant *et al.*, 2007). In contrast, genetic diversity was comparable between salmonellae from asymptomatic human carriers and non-carrier humans plus animals (Chapter 4). Similarly, overall genotypic diversity was comparable between Typhimurium DT104 isolates from asymptomatic and disease associated animal populations (Perron *et al.*, 2007). Reptiles are a natural reservoir for *Salmonella* and carry them asymptomatically in the intestinal tract (Ebani *et al.*, 2005; Geue and Loschner, 2002; Woodward *et al.*, 1997). In agreement with previous observations, similar levels of genotypic and genetic diversity were observed between salmonellae from reptiles and human. However, isolates from NhWBA were less diverse (Table 5.2). This observation is consistent with a previous study

where bovine isolates were found to be less diverse than human isolates (Alcaine *et al.*, 2006). The lower diversity possibly indicates that prevalent genotypes of generalist serovars among NhWBA eliminate others by competitive inhibition. However, competitive inhibition has not yet been described among *Salmonella* isolates. Alternatively, the host environment may provide a selective advantage for particular genotypes in NhWBA hosts. For example, antibiotics are often used to supplement animal feeds (Threlfall *et al.*, 2000; Angulo *et al.*, 2004), which will select for isolates with resistance to those particular antimicrobials, possibly resulting in limited nucleotide and genotypic diversity among animal isolates.

5.4.2 Clonal diversity and host association

Common and host-specific STs were observed within most serovars (Fig. 5.2, Fig 5.3 and Table 5.4). However, most of the specific STs belonged to the same groups as isolates from other hosts (Table 5.5). Only two groups plus four singletons were restricted to isolates from reptiles whereas four groups plus six singletons were specific to human isolates and one group plus two singletons were specific to NhWBA isolates (Fig. 5.2, Table 5.6). The proportion of reptile isolates in host specific groups and singletons was significantly higher than for human or NhWBA isolates. Therefore, host specificity is more frequently exhibited by reptile isolates than isolates from other hosts.

22% (30/137) of all reptile isolates belonged to serovar Oranienburg and most of these were in two host-specific groups. Only five Oranienburg isolates were available from other hosts, all from humans. These isolates shared group 41 with ten Oranienburg isolates from reptiles. The proportion of reptile isolates in host specific groups or singletons (5.6%) was only slightly higher than that of human (3.0%) or NhWBA isolates (1.5%), when Oranienburg isolates were excluded. Therefore, more isolates of serovar Oranienburg from non-reptile hosts are required to be analyzed to determine whether reptile isolates of this serovar are host associated.

None of the host categories represented genetically distinct populations as suggested by the UniFrac significance tests and the P-tests (Table 5.7). This observation might indicate that most genotypes can cross-infect individuals from different host categories. Approximately half of the STs observed among isolates from reptiles as well as NhWBA were shared with human isolates (Fig. 5.3A). This observation suggests that both reptiles and NhWBA are potential reservoirs for *Salmonella* infection in human.

As expected for host adapted serovars, human and cattle isolates of serovar Dublin were identical as were human and swine isolates of serovar Choleraesuis (Fig 3.2, Table 5.5). This observation supports an argument that human infection by serovars Dublin and Choleraesuis probably originate from cattle and pigs, respectively (Kingsley and Baumler, 2000; Uzzau *et al.*, 2000). Most isolates of a host adapted phage type of generalist serovars were confined to a single ST. The only exception was one Typhimurium DT99 isolate whose ST differed from the remaining DT99 isolates. But unlike the host-restricted serovars Paratyphi A, Paratyphi C and Typhi where all isolates were in serovar- and host-specific groups, none of the STs was host specific among host adapted variants of generalist serovars. Enteritidis PT4 and PT11 which are associated to poultry and hedgehogs, respectively, were isolated from other hosts as well. Although Typhimurium DT2 and DT99 are considered to be restricted to pigeons, a reduced ability of isolates of these phage types in murine organs has been reported (Andrews-Polymenis *et al.*, 2004). Since these phage types can infect individuals of other host categories, they do not seem to reflect host specific STs. Furthermore, host adaptation usually involve insertions, deletions and genomic rearrangements (Wu *et al.*, 2005; Helm *et al.*, 2004) that do not necessarily have any impact on housekeeping genes. Genomic rearrangements at *rrn* operon, insertions that potentially contribute to the increased virulence in certain hosts, and genome degradation have been characterized within serovars Typhi and Paratyphi A (Liu *et al.*, 2006; McClelland *et al.*, 2004). The genomes of Typhimurium phage types DT2 and DT99 were similar to Typhimurium LT2 and the only differences were Fels-1 and Fels-2 prophages that were absent in DT2 and DT99 (Andrews-Polymenis *et al.*, 2004). However, these differences did not correlate with the host association when isolates of different phage types were tested (Andrews-Polymenis *et al.*, 2004).

6. Reference

Aarts, H.J., van Lith, L. A., and Keijer, J. (1998) High-resolution genotyping of *Salmonella* strains by AFLP-fingerprinting. *Lett. Appl. Microbiol.* **26**: 131-135.

Achtman, M. (2002) A phylogenetic perspective on molecular epidemiology. In *Molecular Medical Microbiology*. Sussman,M. (ed). London: Academic Press, pp. 485-509.

Achtman, M., Mercer, A., Kusecek, B., Pohl, A., Heuzenroeder, M., Aaronson, W., Sutton, A., and Silver, R. P. (1983) Six widespread bacterial clones among *Escherichia coli* K1 isolates. *Infect. Immun.* **39**: 315-335.

Alcaine, S.D., Soyer, Y., Warnick, L. D., Su, W. L., Sukhnanand, S., Richards, J., Fortes, E. D., McDonough, P., Root, T. P., Dumas, N. B., Grohn, Y., and Wiedmann, M. (2006) Multilocus sequence typing supports the hypothesis that cow- and human-associated *Salmonella* isolates represent distinct and overlapping populations. *Appl. Environ. Microbiol.* **72**: 7575-7585.

Alcaine, S.D., Sukhnanand, S. S., Warnick, L. D., Su, W. L., McGann, P., McDonough, P., and Wiedmann, M. (2005) Ceftiofur-resistant *Salmonella* strains isolated from dairy farms represent multiple widely distributed subtypes that evolved by independent horizontal gene transfer. *Antimicrob. Agents Chemother.* **49**: 4061-4067.

Andrews-Polymenis, H.L., Rabsch, W., Porwollik, S., McClelland, M., Rosetti, C., Adams, L. G., and Baumler, A. J. (2004) Host restriction of *Salmonella enterica* serotype Typhimurium pigeon isolates does not correlate with loss of discrete genes. *J. Bacteriol.* **186**: 2619-2628.

Angulo, F.J., Baker, N. L., Olsen, S. J., Anderson, A., and Barrett, T. J. (2004) Antimicrobial use in agriculture: controlling the transfer of antimicrobial resistance to humans. *Semin. Pediatr. Infect. Dis.* **15**: 78-85.

Angulo, F.J., Johnson, K. R., Tauxe, R. V., and Cohen, M. L. (2000) Origins and consequences of antimicrobial-resistant nontyphoidal *Salmonella*: implications for the use of fluoroquinolones in food animals. *Microb. Drug Resist.* **6**: 77-83.

Animal Disease Diagnostic Laboratory. *Salmonella newport* - An emerging disease in dairy cattle. Animal Disease Diagnostic Laboratory . 2004. http://www.addl.purdue.edu/newsletters/2004/summer/salmnewp.htm

Baker, L., Brown, T., Maiden, M. C., and Drobniewski, F. (2004) Silent nucleotide polymorphisms and a phylogeny for *Mycobacterium tuberculosis*. *Emerg. Infect. Dis.* **10**: 1568-1577.

Baker, S., Hardy, J., Sanderson, K. E., Quail, M., Goodhead, I., Kingsley, R. A., Parkhill, J., Stocker, B., and Dougan, G. (2007) A novel linear plasmid mediates flagellar variation in *Salmonella* Typhi. *PLoS. Pathog.* **3**: e59.

Baumler, A.J., Tsolis, R. M., Ficht, T. A., and Adams, L. G. (1998) Evolution of host adaptation in *Salmonella enterica*. *Infect. Immun.* **66**: 4579-4587.

Beltran, P., Musser, J. M., Helmuth, R., Farmer, J. J., III, Frerichs, W. M., Wachsmuth, I. K., Ferris, K., McWhorter, A. C., Wells, J. G., Cravioto, A., and Selander, R. K. (1988) Toward a population genetic analysis of *Salmonella*: Genetic diversity and relationships among strain of serotypes *S. choleraesuis, S. derby, S. dublin, S. enteritidis, S. heidelberg, S. infantis, S. newport, and S. typhimurium*. *Proc. Natl. Acad. Sci. (U. S. A.)* **85**: 7753-7757.

Beltran, P., Plock, S. A., Smith, N. H., Whittam, T. S., Old, D. C., and Selander, R. K. (1991) Reference collection of strains of the *Salmonella typhimurium* complex from natural populations. *J. Gen. Microbiol.* **137**: 601-606.

Bouchet, V., Huot, H., and Goldstein, R. (2008) Molecular genetic basis of ribotyping. *Clin. Microbiol. Rev.* **21**: 262-73, table.

Boyd, E.F., Nelson, K., Wang, F.-S., Whittam, T. S., and Selander, R. K. (1994) Molecular genetic basis of allelic polymorphism in malate dehydrogenase (*mdh*) in natural populations of *Escherichia coli* and *Salmonella enterica*. *Proc. Natl. Acad. Sci. (U. S. A.)* **91**: 1280-1284.

Boyd, E.F., Porwollik, S., Blackmer, F., and McClelland, M. (2003) Differences in gene content among *Salmonella enterica* serovar Typhi isolates. *J. Clin. Microbiol.* **41**: 3823-3828.

Boyd, E.F., Wang, F. S., Beltran, P., Plock, S. A., Nelson, K., and Selander, R. K. (1993) *Salmonella* reference collection B (SARB): strains of 37 serovars of subspecies I. *J. Gen. Microbiol.* **139**: 1125-1132.

Boyd, E.F., Wang, F. S., Whittam, T. S., and Selander, R. K. (1996) Molecular genetic relationships of the Salmonellae. *Appl. Environ. Microbiol.* **62**: 804-808.

Brenner, F.W., Villar, R. G., Angulo, F. J., Tauxe, R., and Swaminathan, B. (2000) *Salmonella* Nomenclature. *J. Clin. Microbiol.* **38**: 2465-2467.

Briones, V., Tellez, S., Goyache, J., Ballesteros, C., del Pilar, L. M., Dominguez, L., and Fernandez-Garayzabal, J. F. (2004) *Salmonella* diversity associated with wild reptiles and amphibians in Spain. *Environ. Microbiol.* **6**: 868-871.

Brown, E.W., Kotewicz, M. L., and Cebula, T. A. (2002) Detection of recombination among *Salmonella enterica* strains using the incongruence length difference test. *Mol. Phylogenet. Evol.* **24**: 102-120.

Brown, E.W., Mammel, M. K., LeClerc, J. E., and Cebula, T. A. (2003) Limited boundaries for extensive horizontal gene transfer among *Salmonella* pathogens. *Proc. Natl. Acad. Sci. (U. S. A.)* **100**: 15676-15681.

Bruen, T.C., Philippe, H., and Bryant, D. (2006) A simple and robust statistical test for detecting the presence of recombination. *Genetics* **172**: 2665-2681.

Bryant, D. and Moulton, V. (2004) Neighbor-net: an agglomerative method for the construction of phylogenetic networks. *Mol. Biol. Evol.* **21**: 255-265.

Carattoli, A., Tosini, F., Giles, W. P., Rupp, M. E., Hinrichs, S. H., Angulo, F. J., Barrett, T. J., and Fey, P. D. (2002) Characterization of plasmids carrying CMY-2 from expanded-spectrum cephalosporin-resistant *Salmonella* strains isolated in the United States between 1996 and 1998. *Antimicrob. Agents Chemother.* **46**: 1269-1272.

Caugant, D.A., Kristiansen, B.-E., Frøholm, L. O., Bøvre, K., and Selander, R. K. (1988) Clonal diversity of *Neisseria meningitidis* from a population of asymptomatic carriers. *Infect. Immun.* **56**: 2060-2068.

Caugant, D.A., Tzanakaki, G., and Kriz, P. (2007) Lessons from meningococcal carriage studies. *FEMS Microbiol. Rev.* **31**: 52-63.

Caygill, C.P., Hill, M. J., Braddick, M., and Sharp, J. C. (1994) Cancer mortality in chronic typhoid and paratyphoid carriers. *Lancet* **343**: 83-84.

Centers for Disease Control and Prevention (2002) Outbreak of Multidrug-Resistant *Salmonella* Newport in United States, January-April 2002. *Morb. Mortal. Wkly. Rep.* **51**: 545-548.

Centers for Disease Control and Prevention. *Salmonella* Surveillance: Annual Summary, 2005. CDC.Salmonella annual summary. 2006. Georgia: US Human Department of Health and Services.

Centers for Disease control and Prevention (2007) Multistate outbreaks of *Salmonella* infections associated with raw tomatoes eaten in restaurants--United States, 2005-2006. *MMWR Morb. Mortal. Wkly. Rep.* **56**: 909-911.

Centers for Disease Control and Prevention (2007a) *Salmonella* Oranienburg infections associated with fruit salad served in health-care facilities--northeastern United States and Canada, 2006. *MMWR Morb. Mortal. Wkly. Rep.* **56**: 1025-1028.

Centers for Disease Control and Prevention (2007b) *Salmonella* Typhimurium infection associated with raw milk and cheese consumption--Pennsylvania, 2007. *MMWR Morb. Mortal. Wkly. Rep.* **56**: 1161-1164.

Centers for Disease Control and Prevention (2007c) Turtle-associated salmonellosis in humans--United States, 2006-2007. *Morb. Mortal. Wkly. Rep.* **56**: 649-652.

Centres for Disease Control and Prevention, USA . 2007. http://www.cdc.gov/pulsenet/index.htm

Chan, K., Baker, S., Kim, C. C., Detweiler, C. S., Dougan, G., and Falkow, S. (2003) Genomic comparison of *Salmonella enterica* serovars and *Salmonella bongori* by use of an *S. enterica* serovar Typhimurium DNA microarray. *J. Bacteriol.* **185**: 553-563.

Chart, H. (2003) *Salmonella*. In *Medical Microbiology*. Greenwood,D., Slack,R.C.B., and Peutherer,J.F. (eds). London, U.K.: Churchill Livingstone, pp. 250-259.

Cheminay, C. and Hensel, M. (2008) Rational design of *Salmonella* recombinant vaccines. *Int. J. Med. Microbiol.* **298**: 87-98.

Cho, S., Boxrud, D. J., Bartkus, J. M., Whittam, T. S., and Saeed, M. (2007) Multiple-locus variable-number tandem repeat analysis of *Salmonella* Enteritidis isolates from human and non-human sources using a single multiplex PCR. *FEMS Microbiol. Lett.* **275**: 16-23.

Cirillo, V.J. (2000) Fever and reform: the typhoid epidemic in the Spanish-American War. *J. Hist Med. Allied Sci.* **55**: 363-397.

Clark, P.F. (1959) Theobald Smith, Student of Disease (1859-1934). *Journal of the History of Medicine and Allied Sciences* **14**: 490-514.

Cobbold, R.N., Rice, D. H., Davis, M. A., Besser, T. E., and Hancock, D. D. (2006) Long-term persistence of multi-drug-resistant *Salmonella enterica* serovar Newport in two dairy herds. *J. Am. Vet. Med. Assoc.* **228**: 585-591.

Coyle, E.F., Palmer, S. R., Ribeiro, C. D., Jones, H. I., Howard, A. J., Ward, L., and Rowe, B. (1988) *Salmonella enteritidis* phage type 4 infection: association with hen's eggs. *Lancet* **2**: 1295-1297.

Crosa, J.H., Brenner, D. J., Ewing, W. H., and Falkow, S. (1973) Molecular relationships among the Salmonelleae. *J. Bacteriol.* **115**: 307-315.

Cunha, B.A. (2004) The death of Alexander the Great: malaria or typhoid fever? *Infect. Dis. Clin. North Am.* **18**: 53-63.

D'Aoust, J.Y. (1991) Pathogenicity of foodborne *Salmonella*. *Int. J. Food Microbiol.* **12**: 17-40.

de Jong, B., Oberg, J., and Svenungsson, B. (2007) Outbreak of salmonellosis in a restaurant in Stockholm, Sweden, September - October 2006. *Euro. Surveill* **12**: E13-E14.

Denny, J., Threlfall, J., Takkinen, J., Löfdahl, S., Westrell, T., Varela, C., Adak, B., Boxall, N., Ethelberg, S., Torpdahl, M., Straetemans, M., and van Pelt, W. (2007) Multinational *Salmonella* Paratyphi B variant Java (*Salmonella* Java) outbreak, August - December 2007. *Euro. Surveill* **12**: E071220.2.

Devasia, R.A., Varma, J. K., Whichard, J., Gettner, S., Cronquist, A. B., Hurd, S., Segler, S., Smith, K., Hoefer, D., Shiferaw, B., Angulo, F. J., and Jones, T. F. (2005) Antimicrobial use and outcomes in patients with multidrug-resistant and pansusceptible *Salmonella* Newport infections, 2002-2003. *Microbial Drug Resistance-Mechanisms Epidemiology and Disease* **11**: 371-377.

Didelot, X., Achtman, M., Parkhill, J., Thomson, N. R., and Falush, D. (2007) A bimodal pattern of relatedness between the *Salmonella* Paratyphi A and Typhi genomes: Convergence or divergence by homologous recombination? *Genome Res.* **17**: 61-68.

Didelot, X. and Falush, D. (2006) Inference of bacterial microevolution using multilocus sequence data. *Genetics* **DOI:10.1534/genetics.106.063305**.

Dingle, K.E., Colles, F. M., Falush, D., and Maiden, M. C. (2005) Sequence typing and comparison of population biology of *Campylobacter coli* and *Campylobacter jejuni*. *J. Clin. Microbiol.* **43**: 340-347.

Dryden, M.S., Keyworth, N., Gabb, R., and Stein, K. (1994) Asymptomatic foodhandlers as the source of nosocomial salmonellosis. *J. Hosp. Infect.* **28**: 195-208.

Duchet-Suchaux, M., Lechopier, P., Marly, J., Bernardet, P., Delaunay, R., and Pardon, P. (1995) Quantification of experimental *Salmonella enteritidis* carrier state in B13 leghorn chicks. *Avian Dis.* **39**: 796-803.

Dunne, E.F., Fey, P. D., Kludt, P., Reporter, R., Mostashari, F., Shillam, P., Wicklund, J., Miller, C., Holland, B., Stamey, K., Barrett, T. J., Rasheed, J. K., Tenover, F. C., Ribot, E. M., and Angulo, F. J. (2000) Emergence of domestically acquired ceftriaxone-resistant *Salmonella* infections associated with AmpC beta-lactamase. *JAMA* **284**: 3151-3156.

Earle, C. (1979) Environment, disease and mortality in early Virginia. *J. Hist Geogr.* **5**: 365-390.

Ebani, V.V., Cerri, D., Fratini, F., Meille, N., Valentini, P., and Andreani, E. (2005) *Salmonella enterica* isolates from faeces of domestic reptiles and a study of their antimicrobial in vitro sensitivity. *Res. Vet. Sci.* **78**: 117-121.

Egorova, S., Timinouni, M., Demartin, M., Granier, S. A., Whichard, J. M., Sangal, V., Fabre, L., Delaune, A., Pardos, M., Millemann, Y., Espie, E., Achtman, M., Grimont, P. A. D., and Weill, F. X. (2008) Ceftriaxone-resistant *Salmonella enterica* serotype Newport, France. *Emerg. Infect. Dis.* **14**: 954-957.

Emberland, K.E., Ethelberg, S., Kuusi, M., Vold, L., Jensvoll, L., Lindstedt, B. A., Nygard, K., Kjelso, C., Torpdahl, M., Sorensen, G., Jensen, T., Lukinmaa, S., Niskanen, T., and Kapperud, G. (2007) Outbreak of *Salmonella* Weltevreden infections in Norway, Denmark and Finland associated with alfalfa sprouts, July-October 2007. *Euro. Surveill* **12**: E071129.

Emberland, K.E., Nygard, K., Heier, B. T., Aavitsland, P., Lassen, J., Stavnes, T. L., and Gondrosen, B. (2006) Outbreak of *Salmonella* Kedougou in Norway associated with salami, April-June 2006. *Euro. Surveill* **11**: E060706.

Enright, M.C., Robinson, D. A., Randle, G., Feil, E. J., Grundmann, H., and Spratt, B. G. (2002) The evolutionary history of methicillin-resistant *Staphylococcus aureus* (MRSA). *Proc. Natl. Acad. Sci. (U. S. A.)* **99**: 7687-7692.

Espie, E., de Valk, H., Vaillant, V., Quelquejeu, N., Le Querrec, F., and Weill, F. X. (2005) An outbreak of multidrug-resistant *Salmonella enterica* serotype Newport infections linked to the consumption of imported horse meat in France. *Epidemiol. Infect.* **133**: 373-376.

Esteban, E., Snipes, K., Hird, D., Kasten, R., and Kinde, H. (1993) Use of ribotyping for characterization of *Salmonella* serotypes. *J. Clin. Microbiol.* **31**: 233-237.

European Centre for Disease Prevention and Control. Enter-net reports. 2008. http://ecdc.europa.eu/Activities/surveillance/ENTER_NET/reports.html

Euzeby, J.P. (1999) Revised *Salmonella* nomenclature: designation of *Salmonella enterica* (*ex* Kauffmann and Edwards 1952) Le Minor and Popoff 1987 sp. nov., nom. rev. as the neotype species of the genus *Salmonella* Lignieres 1900 (Approved Lists 1980), rejection of the name *Salmonella choleraesuis* (Smith 1894) Weldin 1927 (Approved Lists 1980), and conservation of the name *Salmonella typhi* (Schroeter 1886) Warren and Scott 1930 (Approved Lists 1980). Request for an Opinion. *Int. J. Syst. Bacteriol.* **49**: 927-930.

Euzéby, J.P. 2007. http://www.bacterio.cict.fr/classificationsz.html#Salmonella

Ezaki, T., Amano, M., Kawamura, Y., and Yabuuchi, E. (2000a) Proposal of *Salmonella paratyphi* sp. nov., nom. rev. and Request for an Opinion to conserve the specific epithet *paratyphi* in the binary combination *Salmonella paratyphi* as *nomen epitheton conservandum* . *Int. J. Syst. Evol. Microbiol.* **50**: 941-944.

Ezaki, T., Kawamura, Y., and Yabuuchi, E. (2000b) Recognition of nomenclatural standing of *Salmonella typhi* (Approved Lists 1980), *Salmonella enteritidis* (Approved Lists 1980) and *Salmonella typhimurium* (Approved Lists 1980), and conservation of the specific epithets *enteritidis* and *typhimurium*. Request for an Opinion. *Int. J. Syst. Evol. Microbiol.* **50**: 945-947.

Fabrega, A., Sanchez-Cespedes, J., Soto, S., and Vila, J. (2008) Quinolone resistance in the food chain. *Int. J. Antimicrob. Agents* **31**: 307-315.

Falush, D., Torpdahl, M., Didelot, X., Conrad, D. F., Wilson, D. J., and Achtman, M. (2006) Mismatch induced speciation in *Salmonella*: model and data. *Philos. Trans. R. Soc. Lond B Biol. Sci.* **361**: 2045-2053.

Feil, E.J., Cooper, J. E., Grundmann, H., Robinson, D. A., Enright, M. C., Berendt, T., Peacock, S. J., Smith, J. M., Murphy, M., Spratt, B. G., Moore, C. E., and Day, N. P. (2003) How clonal is *Staphylococcus aureus*? *J. Bacteriol.* **185**: 3307-3316.

Feil, E.J., Holmes, E. C., Bessen, D. E., Chan, M.-S., Day, N. P. J., Enright, M. C., Goldstein, R., Hood, D. W., Kalia, A., Moore, C. E., Zhou, J., and Spratt, B. G. (2001) Recombination within natural populations of pathogenic bacteria: short-term empirical estimates and long-term phylogenetic comparisons. *Proc. Natl. Acad. Sci. (U. S. A.)* **98**: 182-187.

Feil, E.J., Li, B. C., Aanensen, D. M., Hanage, W. P., and Spratt, B. G. (2004) eBURST: Inferring patterns of evolutionary descent among clusters of related bacterial genotypes from Multilocus Sequence Typing data. *J. Bacteriol.* **186**: 1518-1530.

Feil, E.J., Maiden, M. C., Achtman, M., and Spratt, B. G. (1999) The relative contribution of recombination and mutation to the divergence of clones of *Neisseria meningitidis*. *Mol. Biol. Evol.* **16**: 1496-1502.

Feil, E.J., Maynard Smith, J., Enright, M. C., and Spratt, B. G. (2000) Estimating recombinational parameters in *Streptococcus pneumoniae* from multilocus sequence typing data. *Genetics* **154**: 1439-1450.

Felix, A. (1951) Laboratory control of the enteric fevers. *Br. Med. Bull.* **7**: 153-162.

Foley, S.L., Zhao, S., and Walker, R. D. (2007) Comparison of molecular typing methods for the differentiation of *Salmonella* foodborne pathogens. *Foodborne. Pathog. Dis.* **4**: 253-276.

Gay, K., Robicsek, A., Strahilevitz, J., Park, C. H., Jacoby, G., Barrett, T. J., Medalla, F., Chiller, T. M., and Hooper, D. C. (2006) Plasmid-mediated quinolone resistance in non-Typhi serotypes of *Salmonella enterica*. *Clin. Infect. Dis.* **43**: 297-304.

Gebreyes, W.A. and Altier, C. (2002) Molecular characterization of multidrug-resistant *Salmonella enterica* subsp. *enterica* serovar Typhimurium isolates from swine. *J. Clin. Microbiol.* **40**: 2813-2822.

Gebreyes, W.A., Altier, C., and Thakur, S. (2006) Molecular epidemiology and diversity of *Salmonella* serovar Typhimurium in pigs using phenotypic and genotypic approaches. *Epidemiol. Infect.* **134**: 187-198.

Gerner-Smidt, P., Hise, K., Kincaid, J., Hunter, S., Rolando, S., Hyytia-Trees, E., Ribot, E. M., and Swaminathan, B. (2006) PulseNet USA: a five-year update. *Foodborne. Pathog. Dis.* **3**: 9-19.

Gershman, M. (1974) A phage typing system for *Salmonella anatum*. *Avian Dis.* **18**: 565-568.

Gershman, M. (1976) Phage typing system for *Salmonella enteritidis*. *Appl. Environ. Microbiol.* **32**: 190-191.

Gershman, M. and Markowsky, G. (1983) Reduced set of phages for typing salmonellae. *J. Clin Microbiol.* **17**: 240-244.

Geue, L. and Loschner, U. (2002) *Salmonella enterica* in reptiles of German and Austrian origin. *Vet. Microbiol.* **84**: 79-91.

Gordon, M.A. (2008) *Salmonella* infections in immunocompromised adults. *J. Infect.* **56**: 413-422.

Gotelli, N.J. and Colwell, R. K. (2001) Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters* **4**: 379-391.

Gupta, A., Fontana, J., Crowe, C., Bolstorff, B., Stout, A., Van Duyne, S., Hoekstra, M. P., Whichard, J. M., Barrett, T. J., and Angulo, F. J. (2003) Emergence of multidrug-resistant *Salmonella enterica* serotype Newport infections resistant to expanded-spectrum cephalosporins in the United States. *J. Infect. Dis.* **188**: 1707-1716.

Gupta, A., My Thanh, N. T., Olsen, S. J., Sivapalasingam, S., My Trinh, T. T., Phuong Lan, N. T., Hoekstra, R. M., Bibb, W., Minh, N. T., Danh, T. P., Cam, P. D., and Mintz, E.

D. (2006) Evaluation of community-based serologic screening for identification of chronic *Salmonella* Typhi carriers in Vietnam. *Int. J. Infect. Dis.* **10**: 309-314.

Hammer, Ø., Harper, D. A. T., and Ryan, P. D. (2001) PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia Electronica* **4**: -9pp.

Hanage, W.P., Fraser, C., and Spratt, B. G. (2006) The impact of homologous recombination on the generation of diversity in bacteria. *J. theor. Biol.* **239**: 210-219.

Harbottle, H., White, D. G., McDermott, P. F., Walker, R. D., and Zhao, S. (2006) Comparison of multilocus sequence typing, pulsed-field gel electrophoresis, and antimicrobial susceptibility typing for characterization of *Salmonella enterica* serotype Newport isolates. *J. Clin. Microbiol.* **44**: 2449-2457.

Hasian, M.A. (2000) Power, medical knowledge, and the rhetorical invention of "Typhoid Mary. *J. Med. Humanit.* **21**: 123-139.

Helm, R.A., Porwollik, S., Stanley, A. E., Maloy, S., McClelland, M., Rabsch, W., and Eisenstark, A. (2004) Pigeon-associated strains of *Salmonella enterica* serovar Typhimurium phage type DT2 have genomic rearrangements at rRNA operons. *Infect. Immun.* **72**: 7338-7341.

Helmuth, R. and Schroeter, A. (1994) Molecular typing methods for *S. enteritidis*. *Int. J. Food Microbiol.* **21**: 69-77.

Hemsworth, S. and Pizer, B. (2006) Pet ownership in immunocompromised children--a review of the literature and survey of existing guidelines. *Eur. J. Oncol. Nurs.* **10**: 117-127.

Herikstad, H., Motarjemi, Y., and Tauxe, R. V. (2002) *Salmonella* surveillance: a global survey of public health serotyping. *Epidemiol. Infect.* **129**: 1-8.

Heyndrickx, M., Pasmans, F., Ducatelle, R., Decostere, A., and Haesebrouck, F. (2005) Recent changes in *Salmonella* nomenclature: the need for clarification. *Vet. J.* **170**: 275-277.

Hickman-Brenner, F.W., Stubbs, A. D., and Farmer, J. J., III (1991) Phage typing of *Salmonella enteritidis* in the United States. *J. Clin Microbiol.* **29**: 2817-2823.

Hilton, A.C., Banks, J. G., and Penn, C. W. (1997) Optimization of RAPD for fingerprinting *Salmonella*. *Lett. Appl. Microbiol.* **24**: 243-248.

Hohmann, E.L. (2001) Nontyphoidal salmonellosis. *Clin. Infect. Dis.* **32**: 263-269.

Honsa, E., Fricke, T., Stephens, A. J., Ko, D., Kong, F., Gilbert, G. L., Huygens, F., and Giffard, P. M. (2008) Assignment of *Streptococcus agalactiae* isolates to clonal complexes using a small set of single nucleotide polymorphisms. *BMC. Microbiol.* **8**: 140.

Hopkins, K.L., Day, M., and Threlfall, E. J. (2008) Plasmid-mediated quinolone resistance in *Salmonella enterica*, United Kingdom. *Emerg. Infect. Dis.* **14**: 340-342.

Hopkins, K.L., Maguire, C., Best, E., Liebana, E., and Threlfall, E. J. (2007) Stability of multiple-locus variable-number tandem repeats in *Salmonella enterica* serovar Typhimurium. *J. Clin. Microbiol.* **45**: 3058-3061.

Huson, D.H. and Bryant, D. (2006) Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* **23**: 254-267.

Ibrahim, A.E. (1969) Bacteriophage typing of Salmonellae. II. New bacteriophage typing scheme. *Appl. Microbiol.* **18**: 748-754.

Iino, T., Komeda, Y., Kutsukake, K., Macnab, R. M., Matsumura, P., Parkinson, J. S., Simon, M. I., and Yamaguchi, S. (1988) New unified nomenclature for the flagellar genes of *Escherichia coli* and *Salmonella typhimurium*. *Microbiol. Rev.* **52**: 533-535.

Jakobsen, I.B. and Easteal, S. (1996) A program for calculating and displaying compatibility matrices as an aid in determining reticulate evolution in molecular sequences. *CABIOS* **12**: 291-295.

Jolley, K.A., Kalmusova, J., Feil, E. J., Gupta, S., Musilek, M., Kriz, P., and Maiden, M. C. (2000) Carried meningococci in the Czech Republic: a diverse recombining population--. *J. Clin. Microbiol.* **38**: 4492-4498.

Kariuki, S., Revathi, G., Kariuki, N., Kiiru, J., Mwituria, J., Muyodi, J., Githinji, J. W., Kagendo, D., Munyalo, A., and Hart, C. A. (2006) Invasive multidrug-resistant non-typhoidal *Salmonella* infections in Africa: zoonotic or anthroponotic transmission? *J. Med. Microbiol.* **55**: 585-591.

Kidgell, C., Reichard, U., Wain, J., Linz, B., Torpdahl, M., Dougan, G., and Achtman, M. (2002) *Salmonella typhi*, the causative agent of typhoid fever, is approximately 50,000 years old. *Infect. Genet. Evol.* **2**: 39-45.

Kim, S.H., Choi, Y. G., Eom, J. W., Oh, T. J., Lee, K. S., Kim, S. H., Lee, E. T., Park, M. S., Oh, H. B., and Lee, B. K. (2007) An Outbreak of *Salmonella enterica* Serovar Othmarschen at a Funeral Service in Guri-si, South Korea. *Jpn. J. Infect. Dis.* **60**: 412-413.

Kingsley, R.A. and Baumler, A. J. (2000) Host adaptation and the emergence of infectious disease: the *Salmonella* paradigm. *Mol. Microbiol.* **36**: 1006-1014.

Knights, H.T. (1969) Death of a Prince Consort. *Med. J. Aust.* **1**: 480.

Kotetishvili, M., Stine, O. C., Kreger, A., Morris, J. J., Jr., and Sulakvelidze, A. (2002) Multilocus sequence typing for characterization of clinical and environmental *Salmonella* strains. *J. Clin. Microbiol.* **40**: 1626-1635.

Krisztalovics, K., Szabo, E., Danielisz, A., Borbas, K., Paszti, J., and Papp, E. N. (2007) Salmonellosis outbreak in connection with the Formula One race, August 2007 in Hungary. *Euro. Surveill* **12**: E070816.

Lam, S. and Roth, J. R. (1983) IS200: a *Salmonella*-specific insertion sequence. *Cell* **34**: 951-960.

Li, J., Nelson, K., McWhorter, A. C., Whittam, T. S., and Selander, R. K. (1994) Recombinational basis of serovar diversity in *Salmonella enterica*. *Proc. Natl. Acad. Sci. (U. S. A.)* **91**: 2552-2556.

Liebana, E. (2002) Molecular tools for epidemiological investigations of *S. enterica* subspecies *enterica* infections. *Res. Vet. Sci.* **72**: 169-175.

Lindstedt, B.A., Heir, E., Gjernes, E., and Kapperud, G. (2003) DNA fingerprinting of *Salmonella enterica* subsp. *enterica* serovar Typhimurium with emphasis on phage type DT104 based on variable number of tandem repeat loci. *J. Clin. Microbiol.* **41**: 1469-1479.

Lindstedt, B.A., Vardund, T., Aas, L., and Kapperud, G. (2004) Multiple-locus variable-number tandem-repeats analysis of *Salmonella enterica* subsp. *enterica* serovar Typhimurium using PCR multiplexing and multicolor capillary electrophoresis. *J. Microbiol. Methods* **59**: 163-172.

Liu, G.R., Liu, W. Q., Johnston, R. N., Sanderson, K. E., Li, S. X., and Liu, S. L. (2006) Genome plasticity and *ori-ter* rebalancing in *Salmonella typhi*. *Mol. Biol. Evol.* **23**: 365-371.

Liu, G.R., Rahn, A., Liu, W. Q., Sanderson, K. E., Johnston, R. N., and Liu, S. L. (2002) The evolving genome of *Salmonella enterica* serovar Pullorum. *J. Bacteriol.* **184**: 2626-2633.

Liu, Y., Lee, M. A., Ooi, E. E., Mavis, Y., Tan, A. L., and Quek, H. H. (2003) Molecular typing of *Salmonella enterica* serovar Typhi isolates from various countries in Asia by a multiplex PCR assay on variable-number tandem repeats. *J. Clin. Microbiol.* **41**: 4388-4394.

Lozupone, C., Hamady, M., and Knight, R. (2006) UniFrac--an online tool for comparing microbial community diversity in a phylogenetic context. *BMC. Bioinformatics.* **7**: 371.

Lozupone, C. and Knight, R. (2005) UniFrac: a new phylogenetic method for comparing microbial communities. *Appl. Environ. Microbiol.* **71**: 8228-8235.

Ly, K.T. and Casanova, J. E. (2007) Mechanisms of *Salmonella* entry into host cells. *Cell Microbiol.* **9**: 2103-2111.

Mahajan, R.K., Khan, S. A., Chandel, D. S., Kumar, N., Hans, C., and Chaudhry, R. (2003) Fatal case of *Salmonella enterica* subsp. *arizonae* gastroenteritis in an infant with microcephaly. *J. Clin. Microbiol.* **41**: 5830-5832.

Maiden, M.C. (2006) Multilocus sequence typing of bacteria. *Annu. Rev. Microbiol.* **60**: 561-588.

Maiden, M.C.J., Bygraves, J. A., Feil, E., Morelli, G., Russell, J. E., Urwin, R., Zhang, Q., Zhou, J., Zurth, K., Caugant, D. A., Feavers, I. M., Achtman, M., and Spratt, B. G. (1998) Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc. Natl. Acad. Sci. (U. S. A.)* **95**: 3140-3145.

Majtan, T., Majtanova, L., Timko, J., and Majtan, V. (2007) Oligonucleotide microarray for molecular characterization and genotyping of *Salmonella* spp. strains. *J. Antimicrob. Chemother.* **60**: 937-946.

Malorny, B., Bunge, C., Guerra, B., Prietz, S., and Helmuth, R. (2007) Molecular characterisation of *Salmonella* strains by an oligonucleotide multiprobe microarray. *Mol. Cell Probes* **21**: 56-65.

Marr, J.S. and Calisher, C. H. (2003) Alexander the Great and West Nile virus encephalitis. *Emerg. Infect. Dis.* **9**: 1599-1603.

Martin, A.P. (2002) Phylogenetic approaches for describing and comparing the diversity of microbial communities. *Appl. Environ. Microbiol.* **68**: 3673-3682.

Maynard Smith, J., Smith, N. H., O'Rourke, M., and Spratt, B. G. (1993) How clonal are bacteria? *Proc. Natl. Acad. Sci. (U. S. A.)* **90**: 4384-4388.

McCain, C.S. and Powell, K. C. (1990) Asymptomatic salmonellosis in healthy adult horses. *J. Vet. Diagn. Invest* **2**: 236-237.

McClelland, M., Sanderson, K. E., Clifton, S. W., Latreille, P., Porwollik, S., Sabo, A., Meyer, R., Bieri, T., Ozersky, P., McLellan, M., Harkins, C. R., Wang, C., Nguyen, C., Berghoff, A., Elliott, G., Kohlberg, S., Strong, C., Du, F., Carter, J., Kremizki, C., Layman, D., Leonard, S., Sun, H., Fulton, L., Nash, W., Miner, T., Minx, P., Delehaunty, K., Fronick, C., Magrini, V., Nhan, M., Warren, W., Florea, L., Spieth, J., and Wilson, R. K. (2004) Comparison of genome degradation in Paratyphi A and Typhi, human-restricted serovars of *Salmonella enterica* that cause typhoid. *Nat. Genet.* **36**: 1268-1274.

Mead, P.S., Slutsker, L., Dietz, V., McCaig, L. F., Bresee, J. S., Shapiro, C., Griffin, P. M., and Tauxe, R. V. (1999) Food-related illness and death in the United States. *Emerg. Infect. Dis.* **5**: 607-625.

Mermin, J., Hoar, B., and Angulo, F. J. (1997) Iguanas and *Salmonella* Marina infection in children: a reflection of the increasing incidence of reptile-associated salmonellosis in the United States. *Pediatrics* **99**: 399-402.

Mermin, J., Hutwagner, L., Vugia, D., Shallow, S., Daily, P., Bender, J., Koehler, J., Marcus, R., and Angulo, F. J. (2004) Reptiles, amphibians, and human *Salmonella* infection: a population-based, case-control study. *Clin. Infect. Dis.* **38 Suppl 3**: S253-S261.

Miko, A., Pries, K., Schroeter, A., and Helmuth, R. (2005) Molecular mechanisms of resistance in multidrug-resistant serovars of *Salmonella enterica* isolated from foods in Germany. *J. Antimicrob. Chemother.* **56**: 1025-1033.

Millemann, Y., Gaubert, S., Remy, D., and Colmin, C. (2000) Evaluation of IS200-PCR and comparison with other molecular markers to trace *Salmonella enterica* subsp. *enterica* serotype Typhimurium bovine isolates from farm to meat. *J. Clin. Microbiol.* **38**: 2204-2209.

Milstone, A.M., Agwu, A. G., and Angulo, F. J. (2006) Alerting pregnant women to the risk of reptile-associated salmonellosis. *Obstet. Gynecol.* **107**: 516-518.

Morgan, O., Milne, L., Kumar, S., Murray, D., Man, W., Georgiou, M., Verlander, N. Q., de Pinna, E., and McEvoy, M. (2007) Outbreak of *Salmonella* Enteritidis phage type 13a: case-control investigation in Hertsmere, United Kingdom. *Euro. Surveill* **12**: E9-10.

Mortimer, P.P. (1999) Mr N the milker, and Dr Koch's concept of the healthy carrier. *Lancet* **353**: 1354-1356.

Mossong, J., Marques, P., Ragimbeau, C., Huberty-Krau, P., Losch, S., Meyer, G., Moris, G., Strottner, C., Rabsch, W., and Schneider, F. (2007) Outbreaks of monophasic *Salmonella enterica* serovar 4,[5],12:i:- in Luxembourg, 2006. *Euro. Surveill* **12**: E11-E12.

Nakadai, A., Kuroki, T., Kato, Y., Suzuki, R., Yamai, S., Yaginuma, C., Shiotani, R., Yamanouchi, A., and Hayashidani, H. (2005) Prevalence of *Salmonella* spp. in pet reptiles in Japan. *J. Vet. Med. Sci.* **67**: 97-101.

Nauerby, B., Pedersen, K., Dietz, H. H., and Madsen, M. (2000) Comparison of Danish isolates of *Salmonella enterica* serovar Enteritidis PT9a and PT11 from hedgehogs (*Erinaceus europaeus*) and humans by plasmid profiling and pulsed-field gel electrophoresis. *J. Clin. Microbiol.* **38**: 3631-3635.

Navarro, A., Eslava, C., Garcia, d. l. T., Leon, L. A., Licona, D., Leon, L., Zarco, L. A., and Cravioto, A. (2007) Common epitopes in LPS of different Enterobacteriaceae are associated with an immune response against *Escherichia coli* O157 in bovine serum samples. *J. Med. Microbiol.* **56**: 1447-1454.

NCCLS (2002) *Performance standards for antimicrobial disk and dilution susceptibility tests for bacteria isolated from animals. Approved standard, 2nd edition. NCCLS document M31-A2* Wayne, USA: NCCLS.

Nelson, K. and Selander, R. K. (1992) Evolutionary genetics of the proline permease gene (*putP*) and the control region of the proline utilization operon in populations of *Salmonella* and *Escherichia coli*. *J. Bacteriol.* **174**: 6886-6895.

Nelson, K. and Selander, R. K. (1994) Intergeneric transfer and recombination of the 6-phosphogluconate dehydrogenase gene (*gnd*) in enteric bacteria. *Proc. Natl. Acad. Sci. (U. S. A.)* **91**: 10227-10231.

Nelson, K., Wang, F. S., Boyd, E. F., and Selander, R. K. (1997) Size and sequence polymorphism in the isocitrate dehydrogenase kinase/phosphatase gene (*aceK*) and flanking regions in *Salmonella enterica* and *Escherichia coli*. *Genetics* **147**: 1509-1520.

Nelson, K., Whittam, T. S., and Selander, R. K. (1991) Nucleotide polymorphism and evolution in the glyceraldehyde-3-phosphate dehydrogenase gene (*gapA*) in natural populations of *Salmonella* and *Escherichia coli*. *Proc. Natl. Acad. Sci. (U. S. A.)* **88**: 6667-6671.

Nielsen, L.R., Schukken, Y. H., Grohn, Y. T., and Ersboll, A. K. (2004) *Salmonella* Dublin infection in dairy cattle: risk factors for becoming a carrier. *Prev. Vet. Med.* **65**: 47-62.

Octavia, S. and Lan, R. (2006) Frequent recombination and low level of clonality within *Salmonella enterica* subspecies I. *Microbiology* **152**: 1099-1108.

Ollinger-Snyder, P. and Matthews, M. E. (1996) Food safety: review and implications for dietitians and dietetic technicians. *J. Am. Diet. Assoc.* **96**: 163-8, 171.

Olsen, S.J., Bleasdale, S. C., Magnano, A. R., Landrigan, C., Holland, B. H., Tauxe, R. V., Mintz, E. D., and Luby, S. (2003) Outbreaks of typhoid fever in the United States, 1960-99. *Epidemiol. Infect.* **130**: 13-21.

Page, R.D.M. and Holmes, E. C. (1998) *Molecular evolution. A phylogenetic approach* Oxford, UK: Blackwell Science Ltd, pp. 1-346.

Palmera-Suarez, R., Garcia, P., Garcia, A., Barrasa, A., and Herrera, D. (2007) *Salmonella* Kottbus outbreak in infants in Gran Canaria (Spain), caused by bottled water, August-November 2006. *Euro. Surveill* **12**: E070712.

Pasmans, F., Blahak, S., Martel, A., and Pantchev, N. (2007) Introducing reptiles into a captive collection: The role of the veterinarian. *Vet. J.*

Paulley, J.W. (1993) The death of Albert Prince Consort: the case against typhoid fever. *Q. J. Med.* **86**: 837-841.

Perez-Losada, M., Browne, E. B., Madsen, A., Wirth, T., Viscidi, R. P., and Crandall, K. A. (2006) Population genetics of microbial pathogens estimated from multilocus sequence typing (MLST) data. *Infect. Genet. Evol.* **6**: 97-112.

Perron, G.G., Quessy, S., Letellier, A., and Bell, G. (2007) Genotypic diversity and antimicrobial resistance in asymptomatic *Salmonella enterica* serotype Typhimurium DT104. *Infect. Genet. Evol.* **7**: 223-228.

Pezzoli, L., Elson, R., Little, C., Fisher, I., Yip, H., Peters, T., Hampton, M., de Pinna, E., Coia, J. E., Mather, H. A., Brown, D. J., Nielsen, E. M., Ethelberg, S., Heck, M., de Jager, C., and Threlfall, J. (2007) International outbreak of *Salmonella* Senftenberg in 2007. *Euro. Surveill* **12**: E070614.

Popoff, M. (2001) *Antigenic formulas of the Salmonella serovars* Paris, France: WHO collaborating centre for reference and research on *Salmonella*, Institut Pasteur, pp. 1-150.

Popoff, M.Y., Bockemuhl, J., and Gheesling, L. L. (2004) Supplement 2002 (no. 46) to the Kauffmann-White scheme. *Res. Microbiol.* **155**: 568-570.

Poppe, C., Martin, L., Muckle, A., Archambault, M., McEwen, S., and Weir, E. (2006) Characterization of antimicrobial resistance of *Salmonella* Newport isolated from animals, the environment, and animal food products in Canada. *Can. J. Vet. Res.* **70**: 105-114.

Porwollik, S., Boyd, E. F., Choy, C., Cheng, P., Florea, L., Proctor, E., and McClelland, M. (2004) Characterization of *Salmonella enterica* subspecies I genovars by use of microarrays. *J. Bacteriol.* **186**: 5883-5898.

Porwollik, S. and McClelland, M. (2003) Lateral gene transfer in *Salmonella*. *Microbes. Infect.* **5**: 977-989.

Porwollik, S., Mei-Yi, W. R., and McClelland, M. (2002) Evolutionary genomics of *Salmonella*: Gene acquisitions revealed by microarray analysis. *Proc. Natl. Acad. Sci. (U. S. A.)* **99**: 8956-8961.

Posada, D. and Crandall, K. A. (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics.* **14**: 817-818.

Rabsch, W., Andrews, H. L., Kingsley, R. A., Prager, R., Tschape, H., Adams, L. G., and Baumler, A. J. (2002) *Salmonella enterica* serotype Typhimurium and its host-adapted variants. *Infect. Immun.* **70**: 2249-2255.

Rabsch, W., Helm, R. A., and Eisenstark, A. (2004) Diversity of phage types among archived cultures of the Demerec collection of *Salmonella enterica* serovar Typhimurium strains. *Appl. Environ. Microbiol.* **70**: 664-669.

Raetz, C.R. and Whitfield, C. (2002) Lipopolysaccharide endotoxins. *Annu. Rev. Biochem.* **71**: 635-700.

Ramisse, V., Houssu, P., Hernandez, E., Denoeud, F., Hilaire, V., Lisanti, O., Ramisse, F., Cavallo, J. D., and Vergnaud, G. (2004) Variable Number of Tandem Repeats in *Salmonella enterica* subsp. *enterica* for Typing Purposes. *J. Clin. Microbiol.* **42**: 5722-5730.

Reeves, M.W., Evins, G. M., Heiba, A. A., Plikaytis, B. D., and Farmer, J. J., III (1989) Clonal nature of *Salmonella typhi* and its genetic relatedness to other salmonellae as shown by multilocus enzyme electrophoresis, and proposal of *Salmonella bongori* comb. nov. *J. Clin. Microbiol.* **27**: 313-320.

Riley, P.Y. and Chomel, B. B. (2005) Hedgehog zoonoses. *Emerg. Infect. Dis.* **11**: 1-5.

Romani, C., Nicoletti, P., Buonomini, M. I., Nastasi, A., and Mammina, C. (2007) Reinterpreting a community outbreak of *Salmonella enterica* serotype Enteritidis in the light of molecular typing. *BMC. Public Health* **7**: 237.

Roumagnac, P., Weill, F.-X., Dolecek, C., Baker, S., Brisse, S., Chinh, N. T., Le, T. A., Acosta, C. J., Farrar, J., Dougan, G., and Achtman, M. (2006) Evolutionary history of *Salmonella* Typhi. *Science* **314**: 1301-1304.

Rozas, J., Sanchez-DelBarrio, J. C., Messeguer, X., and Rozas, R. (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics.* **19**: 2496-2497.

Rundlof, T., Weintraub, A., and Widmalm, G. (1998) Structural determination of the O-antigenic polysaccharide from *Escherichia coli* O35 and cross-reactivity to *Salmonella arizonae* O62. *Eur. J. Biochem.* **258**: 139-143.

Sadeyen, J.R., Trotereau, J., Protais, J., Beaumont, C., Sellier, N., Salvat, G., Velge, P., and Lalmanach, A. C. (2006) *Salmonella* carrier-state in hens: study of host resistance by a gene expression approach. *Microbes. Infect.* **8**: 1308-1314.

Schaberg, D.R., Tompkins, L. S., and Falkow, S. (1981) Use of agarose gel electrophoresis of plasmid deoxyribonucleic acid to fingerprint Gram-negative bacilli. *J. Clin. Microbiol.* **13**: 1105-1108.

Schmieger, H. (1999) Molecular survey of the *Salmonella* phage typing system of Anderson. *J. Bacteriol.* **181**: 1630-1635.

Schwartz, D.C. and Cantor, C. R. (1984) Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. *Cell 1984. May.* **37**: 67-75.

Selander, R.K., Beltran, P., Smith, N. H., Barker, R. M., Crichton, P. B., Old, D. C., Musser, J. M., and Whittam, T. S. (1990a) Genetic population structure, clonal phylogeny, and pathogenicity of *Salmonella* paratyphi B. *Infect. Immun.* **58**: 1891-1901.

Selander, R.K., Beltran, P., Smith, N. H., Helmuth, R., Rubin, F. A., Kopecko, D. J., Ferris, K., Tall, B. T., Cravioto, A., and Musser, J. M. (1990b) Evolutionary genetic relationships of clones of *Salmonella* serovars that cause human typhoid and other enteric fevers. *Infect. Immun.* **58**: 2262-2275.

Selander, R.K., Caugant, D. A., Ochman, H., Musser, J. M., Gilmour, M. N., and Whittam, T. S. (1986) Methods of multilocus enzyme electrophoresis for bacterial population genetics and systematics. *Applied. Env. Microbiol.* **51**: 873-884.

Selander, R.K., Li, J., Boyd, E. F., Wang, F.-S., and Nelson, K. (1994) DNA sequence analysis of the genetic structure of populations of *Salmonella enterica* and *Escherichia coli*. In *Bacterial systematics and diversity*. Priest,F.G., Ramos-Cormenzana,A., and Tindall,R. (eds). New York: Plenum.

Selander, R.K., Musser, J. M., Caugant, D. A., Gilmour, M. N., and Whittam, T. S. (1987) Population genetics of pathogenic bacteria. *Microb. Pathog.* **3**: 1-7.

Selander, R.K., Smith, N. H., Li, J., Beltran, P., Ferris, K. E., Kopecko, D. J., and Rubin, F. A. (1992) Molecular evolutionary genetics of the cattle-adapted serovar *Salmonella dublin*. *J. Bacteriol.* **174**: 3587-3592.

Shangkuan, Y.H. and Lin, H. C. (1998) Application of random amplified polymorphic DNA analysis to differentiate strains of *Salmonella typhi* and other *Salmonella* species. *J. Appl. Microbiol.* **85**: 693-702.

Shelobolina, E.S., Sullivan, S. A., O'Neill, K. R., Nevin, K. P., and Lovley, D. R. (2004) Isolation, characterization, and U(VI)-reducing potential of a facultatively anaerobic, acid-resistant Bacterium from Low-pH, nitrate- and U(VI)-contaminated subsurface sediment and description of *Salmonella subterranea* sp. nov. *Appl. Environ. Microbiol.* **70**: 2959-2965.

Silverman, M., Zieg, J., Hilmen, M., and Simon, M. (1979) Phase variation in *Salmonella*: genetic analysis of a recombinational switch. *Proc. Natl. Acad. Sci. (U. S. A.)* **76**: 391-395.

Simon, M., Zieg, J., Silverman, M., Mandel, G., and Doolittle, R. (1980) Phase variation: evolution of a controlling element. *Science* **209**: 1370-1374.

Sirinavin, S., Pokawattana, L., and Bangtrakulnondh, A. (2004) Duration of nontyphoidal *Salmonella* carriage in asymptomatic adults. *Clin. Infect. Dis.* **38**: 1644-1645.

Spratt, B.G. (2004) Exploring the concept of clonality in bacteria. *Methods Mol. Biol.* **266**: 323-352.

Srinivasan, A. and McSorley, S. J. (2006) Activation of *Salmonella*-specific immune responses in the intestinal mucosa. *Arch. Immunol. Ther. Exp. (Warsz.)* **54**: 25-31.

Stanley, J., Jones, C. S., and Threlfall, E. J. (1991) Evolutionary lines among *Salmonella enteritidis* phage types are identified by insertion sequence IS*200* distribution. *FEMS Microbiol. Lett.* **82**: 83-9.

Sukhnanand, S., Alcaine, S., Warnick, L. D., Su, W.-L., Hof, J., Craver, M. P. J., McDonough, P., Boor, K. J., and Wiedmann, M. (2005) DNA sequence-based subtyping and evolutionary analysis of selected *Salmonella enterica* serotypes. *J. Clin. Microbiol.* **43**: 3688-3698.

Swofford, D.L. (1998) PAUP*:Phylogenetic analysis using parsimony and other methods. Version 4.0 beta. Sunderland, Mass.: Sinauer Associates.

Tamura, K., Dudley, J., Nei, M., and Kumar, S. (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0. *Mol. Biol. Evol.*

Taylor, D.N., Wachsmuth, I. K., Shangkuan, Y. H., Schmidt, E. V., Barrett, T. J., Schrader, J. S., Scherach, C. S., McGee, H. B., Feldman, R. A., and Brenner, D. J. (1982) Salmonellosis associated with marijuana: a multistate outbreak traced by plasmid fingerprinting. *N. Engl. J. Med.* **306**: 1249-1253.

The National Food Institute. A global *Salmonella* surveillance and laboratory support project of the World Health Organization. Hendriksen, R. S. and Larsen, J. H. 2008. http://www.antimicrobialresistance.dk/232-169-215-protocols.htm

Thomas, A.D., Forbes-Faulkner, J. C., Speare, R., and Murray, C. (2001) Salmonelliasis in wildlife from Queensland. *J. Wildl. Dis.* **37**: 229-238.

Thong, K.L., Cheong, Y. M., Puthucheary, S., Koh, C. L., and Pang, T. (1994) Epidemiologic analysis of sporadic *Salmonella typhi* isolates and those from outbreaks by pulsed-field gel electrophoresis. *J. Clin. Microbiol.* **32**: 1135-1141.

Thorsteinsdottir, T.R., Kristinsson, K. G., and Gunnarsson, E. (2007) Antimicrobial resistance and serotype distribution among *Salmonella* spp. in pigs and poultry in Iceland, 2001-2005. *Microb. Drug Resist.* **13**: 295-300.

Threlfall, E.J., Fisher, I. S., Berghold, C., Gerner-Smidt, P., Tschape, H., Cormican, M., Luzzi, I., Schnieder, F., Wannet, W., Machado, J., and Edwards, G. (2003) Antimicrobial drug resistance in isolates of *Salmonella enterica* from cases of salmonellosis in humans in Europe in 2000: results of international multi-centre surveillance. *Euro. Surveill* **8**: 41-45.

Threlfall, E.J., Hall, M. L., Ward, L. R., and Rowe, B. (1992) Plasmid profiles demonstrate that an upsurge in *Salmonella berta* in humans in England and Wales is associated with imported poultry meat. *Eur. J. Epidemiol.* **8**: 27-33.

Threlfall, E.J. and Peters, T. (2001) Molecular epidemiological analysis of *Salmonella enterica* serotype Derby infections in Hong Kong. *J. Infect.* **42**: 85-86.

Threlfall, E.J., Rowe, B., and Ward, L. R. (1989) Subdivision of *Salmonella enteritidis* phage types by plasmid profile typing. *Epidemiol. Infect.* **102**: 459-465.

Threlfall, E.J., Ward, L. R., Frost, J. A., and Willshaw, G. A. (2000) The emergence and spread of antibiotic resistance in food-borne bacteria. *Int. J. Food Microbiol.* **62**: 1-5.

Tindall, B.J., Grimont, P. A., Garrity, G. M., and Euzeby, J. P. (2005) Nomenclature and taxonomy of the genus *Salmonella*. *Int. J. Syst. Evol. Microbiol.* **55**: 521-524.

Torpdahl, M. and Ahrens, P. (2004) Population structure of *Salmonella* investigated by amplified fragment length polymorphism. *J. Appl. Microbiol.* **97**: 566-573.

Torpdahl, M., Skov, M. N., Sandvang, D., and Baggesen, D. L. (2005) Genotypic characterization of *Salmonella* by multilocus sequence typing, pulsed-field gel electrophoresis and amplified fragment length polymorphism. *J. Microbiol. Methods* **63**: 173-184.

Turner, K.M., Hanage, W. P., Fraser, C., Connor, T. R., and Spratt, B. G. (2007) Assessing the reliability of eBURST using simulated populations with known ancestry. *BMC. Microbiol.* **7**: 30.

Uzzau, S., Brown, D. J., Wallis, T., Rubino, S., Leori, G., Bernard, S., Casadesus, J., Platt, D. J., and Olsen, J. E. (2000) Host adapted serotypes of *Salmonella enterica*. *Epidemiol. Infect.* **125**: 229-255.

Van Belkum, A. (1999) Short sequence repeats in microbial pathogenesis and evolution. *Cell Mol. Life Sci.* **56**: 729-734.

Van Belkum, A., Scherer, S., van Alphen, L., and Verbrugh, H. (1998) Short-sequence DNA repeats in prokaryotic genomes. *Microbiol. Mol. Biol. Rev.* **62**: 275-293.

van Duijkeren, E., Wannet, W. J., Houwers, D. J., and van Pelt, W. (2002) Serotype and phage type distribution of *Salmonella* strains isolated from humans, cattle, pigs, and chickens in the Netherlands from 1984 to 2001. *J. Clin. Microbiol.* **40**: 3980-3985.

van Lier, E.A. and Havelaar, A. H. The burden of infectious diseases in Europe: a pilot Study. RIVM Report 215011001/2007, 1-48. 2007. Bilthoven: National Institute for Public Health and the Environment (RIVM), The Netherlands.

Voetsch, A.C., Van Gilder, T. J., Angulo, F. J., Farley, M. M., Shallow, S., Marcus, R., Cieslak, P. R., Deneen, V. C., and Tauxe, R. V. (2004) FoodNet estimate of the burden of illness caused by nontyphoidal *Salmonella* infections in the United States. *Clin. Infect. Dis.* **38 Suppl 3**: S127-S134.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de, L. T., Hornes, M., Frijters, A., Pot, J., Peleman, J., and Kuiper, M. (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.* **23**: 4407-4414.

Wachsmuth, I.K., Kiehlbauch, J. A., Bopp, C. A., Cameron, D. N., Strockbine, N. A., Wells, J. G., and Blake, P. A. (1991) The use of plasmid profiles and nucleic acid probes in epidemiologic investigations of foodborne, diarrheal diseases. *Int. J. Food Microbiol.* **12**: 77-89.

Wang, F.S., Whittam, T. S., and Selander, R. K. (1997) Evolutionary genetics of the isocitrate dehydrogenase gene (*icd*) in *Escherichia coli* and *Salmonella enterica*. *J. Bacteriol.* **179**: 6551-6559.

Weill, F.X., Guesnier, F., Guibert, V., Timinouni, M., Demartin, M., Polomack, L., and Grimont, P. A. (2006) Multidrug resistance in *Salmonella enterica* serotype Typhimurium from humans in France (1993 to 2003). *J. Clin. Microbiol.* **44**: 700-708.

Werner, S., Boman, K., Einemo, I., Erntell, M., de Jong, B., Lindqvist, A., Lofdahl, M., Lofdahl, S., Meeuwisse, A., Ohlen, G., Olsson, M., Stamer, U., Sellstrom, E., and Andersson, Y. (2007) Outbreak of *Salmonella* Stanley in Sweden associated with alfalfa sprouts, July-August 2007. *Euro. Surveill* **12**: E071018.

Williams, A.N. and Arnott, R. (2004) A stone at the Siege of Cyropolis and the death of Alexander the Great. *J. Hist Neurosci.* **13**: 130-137.

Willis, C., Wilson, T., Greenwood, M., and Ward, L. (2002) Pet reptiles associated with a case of salmonellosis in an infant were carrying multiple strains of *Salmonella*. *J. Clin. Microbiol.* **40**: 4802-4803.

Winokur, P.L., Brueggemann, A., DeSalvo, D. L., Hoffmann, L., Apley, M. D., Uhlenhopp, E. K., Pfaller, M. A., and Doern, G. V. (2000) Animal and human multidrug-resistant, cephalosporin-resistant *Salmonella* isolates expressing a plasmid-mediated CMY-2 AmpC beta-lactamase. *Antimicrob. Agents Chemother.* **44**: 2777-2783.

Woodward, D.L., Khakhria, R., and Johnson, W. M. (1997) Human salmonellosis associated with exotic pets. *J. Clin. Microbiol.* **35**: 2786-2790.

World Health Organization. Drug-resistant *Salmonella*. Food Safety Department, WHO . 2005. http://www.who.int/mediacentre/factsheets/fs139/en/index.html

Worley, M.J., Nieman, G. S., Geddes, K., and Heffron, F. (2006) *Salmonella typhimurium* disseminates within its host by manipulating the motility of infected cells. *Proc. Natl. Acad. Sci. U. S. A* **103**: 17915-17920.

Wu, K.Y., Liu, G. R., Liu, W. Q., Wang, A. Q., Zhan, S., Sanderson, K. E., Johnston, R. N., and Liu, S. L. (2005) The genome of *Salmonella enterica* serovar Gallinarum: distinct insertions/deletions and rare rearrangements. *J. Bacteriol.* **187**: 4720-4727.

Xiang, S.H., Hobbs, M., and Reeves, P. R. (1994) Molecular analysis of the rfb gene cluster of a group D2 *Salmonella enterica* strain: evidence for its origin from an insertion sequence-mediated recombination event between group E and D1 strains. *J. Bacteriol.* **176**: 4357-4365.

Yabuuchi, E. and Ezaki, T. (2000) Arguments against the replacement of type species of the genus *Salmonella* from *Salmonella choleraesuis* to '*Salmonella enterica*' and the creation of the term 'neotype species', and for conservation of *Salmonella choleraesuis*. *Int. J. Syst. Evol. Microbiol.* **50 Pt 4**: 1693-1694.

Yan, S.S., Pendrak, M. L., Abela-Ridder, B., Punderson, J. W., Fedorko, D. P., and Foley, S. L. (2003) An overview of *Salmonella* typing: Public health perspectives. *Clinical and Applied Immunology Reviews*: 189-204.

Yang, Z. (2007) PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**: 1586-1591.

Yang, Z., Nielsen, R., Goldman, N., and Pedersen, A. M. (2000) Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* **155**: 431-449.

Yazdankhah, S.P., Kriz, P., Tzanakaki, G., Kremastinou, J., Kalmusova, J., Musilek, M., Alvestad, T., Jolley, K. A., Wilson, D. J., McCarthy, N. D., Caugant, D. A., and Maiden, M. C. (2004) Distribution of serogroups and genotypes among disease-associated and carried isolates of *Neisseria meningitidis* from the Czech Republic, Greece, and Norway. *J. Clin. Microbiol.* **42**: 5146-5153.

Yoshida, C., Franklin, K., Konczy, P., McQuiston, J. R., Fields, P. I., Nash, J. H., Taboada, E. N., and Rahn, K. (2007) Methodologies towards the development of an oligonucleotide microarray for determination of *Salmonella* serotypes. *J. Microbiol. Methods* **70**: 261-271.

Zhao, S., White, D. G., McDermott, P. F., Friedman, S., English, L., Ayers, S., Meng, J., Maurer, J. J., Holland, R., and Walker, R. D. (2001) Identification and expression of cephamycinase bla(CMY) genes in *Escherichia coli* and *Salmonella* isolates from food animals and ground meat. *Antimicrob. Agents Chemother.* **45**: 3647-3650.

7. Publications

1. Sangal, V., Bernhoeft, S., Curtiss-III, R., Achtman, M. *(in prearation)* Genetic relationships between *Salmonella* from human carriers and from disease.

2. Sangal, V., Harbottle, H., Mazzoni, C., Roumagnac, P., Helmuth, R., Guerra-Román, B., Weill, F-X., Rabsch, W., Achtman, M. (*in preperation)* Evolution and population structure of *Salmonella enterica* serovar Newport.

3. Egorova, S., Timinouni, M., Demartin, M., Granier, S.A., Whichard, J.M., Sangal, V., Fabre, L., Delauné, A., Pardos, M., Millemann, Y., Espié, E., Achtman, M., Grimont, P.A.D., Weill, F-X. (2008) Ceftriaxone-resistant *Salmonella enterica* serotype Newport, France. *Emerging Infectious Diseases* 14(6):954-7.

4. Bashyam M.D., Bashyam, L., Savithri, G.R., Gopikrishna, M., Sangal, V., Devi, A.R. (2004) Molecular genetic analyses of *β*-thalassemia in South India reveals rare mutations in the *β*-globin gene. *Journal of Human Genetics* 49(8):408-13.

5. Ahmed, N., Alam, M., Rao, K.R., Kauser, F., Kumar, N.A., Qazi, N.N., Sangal, V., Sharma, V.D., Das, R., Katoch, V.M., Murthy, K.J., Suneetha, S., Sharma, S.K., Sechi, L.A., Gilman, R.H., Hasnain, S.E. (2004) Molecular genotyping of a large, multicentric collection of tubercle bacilli indicates geographical partitioning of strain variation and has implications for global epidemiology of *Mycobacterium tuberculosis*. *Journal of Clinical Microbiology* 42(7):3240-7.

6. Ahmed, N., Caviedes, L., Alam, M., Rao, K.R., Sangal, V., Sheen, P., Gilman, R.H., Hasnain, S.E. (2003) Distinctiveness of Mycobacterium tuberculosis genotypes from human immunodeficiency virus type 1-seropositive and -seronegative patients in Lima, Peru. *Journal of Clinical Microbiology* 41(4):1712-6.

8. SELBSTÄNDIGKEITSERKLÄRUNG

Hiermit erkläre ich, die Dissertation selbständig und nur unter Verwendung der angegebenen Hilfen und Hilfsmittel angefertigt habe.
Ich habe mich anderwärts nicht um einen Doktorgrad beworben und besitze einen entsprechenden Doktorgrad nicht.
Ich erkläre die Kenntnisnahme der dem Verfahren zugrunde liegenden Promotionsordnung der Mathematisch-Naturwissenschaftlichen Fakultät I der Humboldt-Universität zu Berlin.

Berlin, den 13th October, 2008 Vartul Sangal